

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2004, 23:30:45 ; Search time 529 Seconds
(without alignments)
2589.979 Million cell updates/sec

Title: US-10-043-649-2
Perfect score: 1351
Sequence: 1 MGSLPSRRKSLPSPSLSSSV.....RESLSFYISLNDKAVSLDDA 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spool/US10043649/runat_16112004_060535_28709/app query.fasta_1.455
-DB=N Geneseq_23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10043649 @CGN 1 1 470 @runat_16112004_060535_28709 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	1351	100.0	786 6	AAL44089 Human mod
2	1351	100.0	786 6	ABQ74343 Human Src
3	1351	100.0	2567 6	AAD43980 Human Src
c 4	1351	100.0	3646 10	ADF90741 Human hep
5	1347	99.7	1183 6	ABK61465 Human CDN
6	1345	99.6	2788 10	ADF82458 Adf82458 Leukaemia

7	1273	94.2	837	3	AAC77202	Aac77202 Human ORF
8	1210.5	89.6	1413	6	ABQ99374	Abq99374 Human cod
9	1200.5	88.9	737	6	AAL44090	Aal444090 Mouse MAR
10	1132.5	83.8	2049	5	AAS74750	Aas74750 DNA encod
11	1032	76.4	1348	6	AAL44087	Aal444087 Mouse mod
12	830	61.4	763	5	ABQ98670	Abq98670 Human ORF
13	731	54.1	603	5	AAS74748	Aas74748 DNA encod
14	647	47.9	864	5	ADL63090	Adl63090 Human ova
15	586	43.4	875	6	ABQ99151	Abq99151 Human ORF
16	492	36.4	3756	8	ABX62975	Abx62975 Human act
17	491	36.3	2109	4	AAS02049	Aas02049 DNA encod
18	491	36.3	2665	6	ABL65189	AbL65189 Lung canc
19	491	36.3	2665	6	ABK83738	Abk83738 Human CDN
20	491	36.3	2665	8	ACC81091	Acc81091 Human Src
21	491	36.3	2665	12	ADL83115	Adl83115 Human PRO
22	491	36.3	2665	12	ADP12833	Adp12833 Reference
23	491	36.3	3452	12	ADQ22520	Adq22520 Human sof
24	452.5	33.5	444	6	ABQ98669	Abq98669 Human ORF
25	374.5	27.7	2298	6	ABK83935	Abk83935 Human CDN
26	374.5	27.7	2298	10	ADD19015	Add19015 Human dis
27	374.5	27.7	2298	10	ADF81586	Adf81586 leukaemia
28	374.5	27.7	2298	11	AD131832	Ad131832 Human CDN
29	374.5	27.7	2298	11	ADN95431	Adn95431 Human BEC
30	374.5	27.7	2298	12	ADL22891	Adl22891 Human MP2
31	374.5	27.7	2298	12	ADN04497	Adn04497 Antipsori
32	374.5	27.7	2298	12	ADP10425	Adp10425 Reference
33	374.5	27.7	4175	10	ADF81585	Adf81585 leukaemia
34	364.5	27.0	1924	12	ADI04091	Adi04091 Human HCK
35	364.5	27.0	1926	6	ABK83940	Abk83940 Human CDN
36	364.5	27.0	1926	12	ADJ71658	Adj71658 Human NOV
37	364.5	27.0	2015	6	ABL66673	AbL66673 lung canc
38	364.5	27.0	2015	6	ABK83939	Abk83939 Human CDN
39	364.5	27.0	2015	10	AAD62155	Aad62155 Human hae
40	364.5	27.0	2015	11	ADI31779	Adi31779 Human CDN
41	364.5	27.0	2015	12	ADL22888	Adl22888 Human MP2
42	364.5	27.0	2015	12	ADP12819	Adp12819 Reference
43	364.5	27.0	2025	11	ADL71038	Adl71038 Gene enco
44	364.5	27.0	2341	10	ADE53793	Ades3793 Human pro
45	362.5	26.8	1911	6	ABK63704	Abk63704 Rat seque

ALIGNMENTS

RESULT 1		AAL44089 standard; cDNA; 786 BP.	
ID	AAL44089		
AC	AAL44089;		
DT	03-OCT-2002	(first entry)	
XX	Human modulator of antigen receptor signalling protein coding sequence.		
DE			
XX			
KW	Human; gene; ss; gene therapy; modulator of antigen receptor signalling;		
KW	MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;		
KW	myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;		
KW	immunosuppression; myeloproliferative disorder; breast cancer.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..786	
FT		/*tag= a	
FT		/product= "Human MARS protein"	
XX			
PN	WO200242452-A2.		
XX			
PD	30-MAY-2002.		
XX			
PF	26-NOV-2001; 2001WO-CA001662.		
XX			
PR	27-NOV-2000; 2000CA-02324663.		
XX			

PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Mcglade JC, Loreto MP;
XX
DR WPI; 2002-566564/60.
DR P-PSDB; AA015457.
XX
PT New isolated modulator of antigen receptor signaling protein or its
PT fragment, useful for treating malignant disorders such as myeloid
PT malignancies, autoimmune disorders and myeloproliferative disorders.
XX
PS Claim 12; Page 75; 110pp; English.
XX
CC The invention comprises the amino acid and coding sequences of modulator
CC of antigen receptor signaling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present CDNA sequence encodes a human MARS protein
XX
SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.36e-110 Length: 786
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AAL44089 (1-786)

QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerVal 20
|||
Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCCCTGTGC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
|||
Db 61 CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGAGAGAGCCACAGCCGTGGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
|||
Db 121 GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
|||
Db 181 ATCGTCTCTGAGGATGAGACTGTGTGACGGTGTCTGTGAAGTCTCAGGCAGAGATAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
|||
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
|||
Db 301 AGGAGAAAGCAGAGAGAACTGCTGTGTACTCTGGGAACCTTGAGGGGCTTCTCATTC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
|||
Db 361 CGGAGAGCCAGACGAGAGAGGCTTACTCTCTGTGAGTCCGCCCTCAGCCGCCCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
|||
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACAATGGCTGGCTGTACATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
|||
Db 481 TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCCCTGTGGACCATTAATCTGAGCTGGCG 540
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
|||

Db 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTGTGACAGAGGCTGGCCGCTCCCT 600
QY 201 GLYLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
|||
Db 601 GGCAAGGATATACCCCTACTGTGACTGTGACAGAGACACCACTCAACTGAAAGAGCTG 660
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
|||
Db 661 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGAGAGTCTCTTCTCAGTGAAGGT 720
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuLysAspGluAlaValSerLeuAspAsp 260
|||
Db 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGATGAT 780
QY 261 Ala 261
|||
Db 781 GCC 783

RESULT 2
ABQ74343
ID ABQ74343 standard; cDNA; 786 BP.
XX
AC ABQ74343;
XX
DT 15-OCT-2002 (first entry)
XX

Human Src-like inhibitory molecule (SLIM) encoding cDNA.

Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;
SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;
modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;
KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;
KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;
KW gene; ss.

XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 1..786
FT /product= "SLIM"
FT /note= "Src-like inhibitory molecule"

XX WO200255707-A2.
XX
XX 18-JUL-2002.
XX
XX 10-JAN-2002; 2002WO-US000718.
XX
XX 10-JAN-2001; 2001US-0260953P.
XX

XX (RIGE-) RIGEL PHARM INC.
XX
XX Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;
PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;
XX WPI; 2002-575432/61.
DR P-PSDB; ABP52187.
XX

XX New src-like inhibitory molecule protein, useful for treating
PT immunodeficiency disorders and inflammatory disorders, comprises N-
PT terminal myristylation sequence, SH2 domain and/or SH3 domain.
XX

PS Claim 3; Fig 2A; 91pp; English.

XX The present sequence encodes the human Src-like inhibitory molecule
CC (SLIM) protein (I). The present invention describes a SLIM protein
CC comprising an N-terminal myristylation sequence, an N-terminal SH2
CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising
CC an N-terminal myristylation sequence and an N-terminal SH2 domain which
CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and
CC anti-HIV activities, and can be used as a modulator of lymphocyte
CC activation, and of ubiquitination of a Cbl target protein, and in gene

CC therapy. (I) is useful for screening a bioactive agent capable of binding
CC to SLIM. (I) is also useful for screening a bioactive agent capable of
CC modulating SLIM binding. (I) or its fragments is useful in the study or
CC in the treatment of conditions which involves this function or
CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or
CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it
CC (II) is useful for modulating leukocyte and/or platelet activation, for
CC modulating antigen receptor-induced signaling and activation in
CC leukocyte and/or platelets and for modulating antigen receptor-induced
CC signaling and activation in lymphocytes and/or mast cells. (I) or (II)
CC is also useful for modulating the basal activity of lymphocytes. (I) or
CC (II) is useful in the treatment of immunodeficiency disorders, such as
CC acquired immunodeficiency syndrome (AIDS), for the prevention and
CC treatment of acute inflammatory disorders, chronic inflammatory
CC disorders, autoimmune disorder and transplant rejection

XX
SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.36e-110 Length: 786
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x ABQ74343 (1-786)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAAGTTTCCCGCAGGTGGCCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCGTCTGTAGGATGAGACTGTGTGACGGTGTCTGTAAGTCTCAAGGCACAGAGATAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCCCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 301 AGGAGAAAGCAGAGAACTGCTGTGTACTGGAACCCCTGAGGGGCTTCCCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGGAGAGCCAGACAGAGAGAGGCTTACTCTCTCTCAGTCCGCCCTCAGCCGCCCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACATATGGCTGTATCATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTAATTGAGCTGGCG 540
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTCTGAGAGGGCTGGCCGCTCCCT 600
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 601 GGCAAGATATACCTTACTGTGACTGTGACAGAGACCACTCAACTGGAAAGAGCTG 660
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240

Db 661 GACAGCTCCCTCTGTTTCTTGAAGCTGCCACAGGAGAGAGTCTTCTTCACTGAGGGT 720
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGATGAT 780
QY 261 Ala 261
Db 781 GCC 783
RESULT 3
AAD43980
ID AAD43980 standard, cDNA; 2567 BP.
XX
AC AAD43980;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.
XX
KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;
KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;
KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;
KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;
KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;
KW multiple sclerosis; asthma; acute respiratory distress syndrome;
KW pulmonary disorder; dermatological; neuroprotective; gene; ss.
KW
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 415..1200
FT /*tag= a
FT /product= "Human SLAP-2"
XX
PN WO200242457-A1.
XX
PD 30-MAY-2002.
XX
PE 20-NOV-2001; 2001WO-US043367.
XX
PR 22-NOV-2000; 2000US-0252545P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;
XX
XX WPI; 2002-463632/49.
DR P-PSDB; AAB26357.
XX
PT Novel substantially purified human SH2/SH3-domain-containing adapter
PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic
PT intervention in immunological and inflammatory disorders and cancer.
XX
PS Claim 2; Fig 1; 85pp; English.
XX
CC The invention relates to a substantially purified human SH2/SH3-domain-
CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-
CC 2). The invention is useful for treating an immune disorder involving
CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is
CC useful for screening for antagonists or inhibitors of the interaction of
CC hSLAP-2 with cellular signaling compounds, for diagnosing, treating or
CC preventing diseases or disorders associated with aberrant or uncontrolled
CC cellular signal transduction, for determining those cellular signaling
CC molecules which associate with hSLAP-2 and which provide critical signals
CC for cell activation, and as effectors in methods to affect T- cell
CC activation. The invention is useful in screening assays to identify and
CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for
CC potential use to treat autoimmune diseases which may be caused by
CC hyperactivated B cells, as well as to treat diseases which may be caused
CC by hyperactivated T cells, in addition to other immune system related
CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,

CC inflammation disorders, diseases and conditions, rheumatoid arthritis, CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's CC and ulcerative colitis), allergies, particularly those involving CC hyperactivity of B-cells and T-cells, or other immune cells, such as CC mast cells or eosinophils, autoimmune diseases such as systemic lupus erythematosus and multiple sclerosis, pulmonary diseases including CC asthma, acute respiratory distress syndrome, and chronic obstructive CC pulmonary disorder, tissue/ organ rejection and cancer. The invention is CC useful in gene therapy. The present sequence is human SLAP-2 cDNA

XX
SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.29e-110 Length: 2567
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AAD43980 (1-2567)

QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerSerVal 20
Db 415 ATGGGAAGTCTGCCAGCAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTGTGC 474
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 475 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGCCACAGCCGTGCCCTG 534
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 535 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 594
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 595 ATCGTCTGTAGGATGGAGACTGCTGTGACCGGTGCTGTGAAGTCTCAGGCAAGAGTAT 654
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 655 AACATCCCCAGCGTCCACGTGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGACC 714
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 715 AGGAGAAAGCAGAGGAACCTGTTGTACTCTGGAAACCTTGAGGGCCTTCCTCATC 774
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 775 CGGAGAGCCAGACCAGAGAGGCTTACTCTCTGTCACTCCGCCCTCAGCCGCCCTGCA 834
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 835 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACATGGCTGGCTGTACATC 894
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 895 TCACCGCGCCTCACCTTCCCTCACTCCAGGCGCCTGGTGACCATTACTCTGAGCTGGCG 954
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 955 GATGACATCTGTGCTGCTACTCAAGAGCCCTGTGTCTGCAAGGGGCTGGCCGCTCCCT 1014
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 1015 GGCAGAGATATACCCCTACTGTTGACTGTGCAGAGACACCACTCAACTGGAAGAGCTTG 1074
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 1075 GACAGCTCCCTCCTGTTTCTGAGACTGCCACAGAGGAGAGTCTCTTCTCAGTGAGGT 1134
QY 241 IleuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 1135 CTCGGGAGTCCCTCAAGTCTTACATCAGCCTGAATGACGAGGCTGTCTCTTTGATGAT 1194

QY 261 Ala 261
Db 1195 GCC 1197

RESULT 4
ADPF0741/c
ID ADPF0741 standard; DNA; 3646 BP.
XX
XX ADF90741;
AC
XX
XX 26-FEB-2004 (first entry)
DT
XX
DE Human hepatic-fibrosis disease marker SEQ ID 203.

XX
XX Hepatic fibrosis; marker; chronic hepatitis; liver cirrhosis;
KW hepatic carcinoma; human; ds.
XX
OS Homo sapiens.
XX
PN JP2003259877-A.
XX
PD 16-SEP-2003.
XX
PF 11-MAR-2002; 2002JP-00065013.
XX
PR 11-MAR-2002; 2002JP-00065013.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.
XX

DR WPI; 2003-821598/77.
XX

PT Hepatic fibrosis disease markers comprising polynucleotides or
PT antibodies, useful for improved diagnosis, screening and developing drugs
PT to treat hepatitis, to control cirrhosis and carcinoma.
XX
PS Claim 1, SEQ ID NO 203; 313pp; Japanese.

XX
CC The present invention relates to hepatic-fibrosis disease markers
CC (ADPF90539-ADPF90871) and related proteins (ADPF90872-ADPF90917). The
CC sequences are useful for detecting and treating hepatic fibrosis caused
CC by alcohol consumption, virus infection, etc., and the associated chronic
CC hepatitis, etc. leading to liver cirrhosis and hepatic carcinoma. The
CC markers allow the cause of hepatic fibrosis to be clarified (diagnostic
CC precision), so more suitable treatments can be developed and given.
XX

SQ Sequence 3646 BP; 782 A; 954 C; 1031 G; 877 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 9.91e-110 Length: 3646
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-043-649-2 (1-261) x ADF90741 (1-3646)

QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerSerVal 20
Db 3202 ATGGGAAGTCTGCCAGCAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTGTGC 3143
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 3142 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGCCACAGCGTGGCCCTG 3083
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 3082 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 3023
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 3022 ATGTTCTTGAGGATGAGACTGGTGAACGGTGTGTCTGAAGTCTCAGGACAGAGATAT 2963

QY 81 AsnIleProSerValHisValAlaIalysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 2962 AACATCCCCAGCGTCCACGTCGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGACC 2903
QY 101 ArgGluIysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 2902 AGGGAGAAAGCAGAGAACTGCTGTGTACTCTGTGAGTCCGCTCAGCGCCCTTCATC 2843
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 2842 CCGGAGAGCCAGACCAGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCGCCCTGACA 2783
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 2782 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACAAATGGCTGGCTGTACATC 2723
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 2722 TCACCGCGCTCACCCTTCCCTCACTCCAGGCCCTGGTGACCAATTACTCTGAGCTGGCG 2663
QY 181 AspAspIleCysCysLeuLeuIysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 2662 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGTCTGACAGAGGGCTGGCCGCTCCCT 2603
QY 201 GlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpIysGluLeu 220
Db 2602 GGCAAGGATATACCCCTACTGCTGACTGTGACAGAGGACCACTCAACTGGAAGAGCTG 2543
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 2542 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTTCTCAGTGAAGGT 2483
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 2482 CTCCGGAGTCCCTCACTCTTACATCAGCCTGAATGACGAGGCTGTCTTTGATGAT 2423
QY 261 Ala 261
Db 2422 GCC 2420
RESULT 5
ABK61465
ID ABK61465 standard; cDNA; 1183 BP.
XX
AC ABK61465;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human cDNA encoding protein NOV13.
XX
KW Human; gene; ss; NOX; gene therapy; cardiomyopathy; atherosclerosis;
KW cell signal processing disorder; metabolic pathway modulation disorder;
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;
KW uterus cancer; immune response; graft-versus-host disease;
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW Albright hereditary osteodystrophy.
XX
OS Homo sapiens.
XX
PN WO200216599-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US026510.
XX
PR 25-AUG-2000; 2000US-0228191P.
PR 08-FEB-2001; 2001US-0267300P.
PR 20-FEB-2001; 2001US-0269961P.
PR 20-MAR-2001; 2001US-0277337P.
XX
PA (CURA-) CURAGEN CORP.

PA (CORT-) COR THERAPEUTICS INC.
XX
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX
DR WPI; 2002-280937/32.
DR P-PSDB; AAU91308.
XX
PT New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
PS Claim 1; Page 98; 263pp; English.
XX
CC The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
CC and antibody against it, are useful for treating or preventing (e.g. by
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC atherosclerosis, a disorder related to cell signal processing and
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC and nucleic acids are also useful for determining the presence of
CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC especially useful in therapeutic or prophylactic applications for
CC disorders associated with aberrant NOVX expression or activity, e.g.
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC cancer), immune response, graft-versus-host disease, acquired
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC congenital heart defects, multiple sclerosis, inflammation or Albright
CC hereditary osteodystrophy and many other diseases listed in the
CC specification. The DNA encoding the protein is useful in gene therapy for
CC treating the conditions. This is also useful in detection assays,
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC for developing a powerful assay system for functional analysis of various
CC human disorders, as well as in diagnostic applications. The present
CC sequence encodes a NOVX protein
XX
SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5:22e-110 Length: 1183
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.70% Indels: 0
DB: 6 Gaps: 0
US-10-043-649-2 (1-261) x ABK61465 (1-1183)
QY 1 MetGlySerLeuProSerArgArgIysSerLeuProSerProSerLeuSerSerVal 20
Db 398 ATGGAAGTCTGCCAGCAGAGAAATAATCTGTGCCAAAGCTTGAGTTCTCTGTGC 457
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40
Db 458 CAAGGCCAGGAGCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 517
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 518 GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCCATTGACC 577
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 578 ATCGTCTTGAGATGAGACTGCTGAGCGGTGCTGTGAAGTCTCAGCAGAGAGTAT 637
QY 81 AsnIleProSerValHisValAlaIalysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 638 AACATCCCCAGCGTCCACGTCGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGAGC 697
QY 101 ArgGluIysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 698 AGGGAGAAAGCAGAGAACTGCTGTGTACTCTGGAACCTGAGGGGCTTCCTCATC 757

Db 1167 GCC 1169

RESULT 7
AAC77202

ID AAC77202 standard; cDNA; 837 BP.
XX AC AAC77202;
XX DT 08-FEB-2001 (first entry)
XX DE Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.

KW Human; open reading frame; ORFX; detection; cyostatic; hepatotropic;
KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticovulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antihydroid;
KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.
OS WO200058473-A2.
PN 05-OCT-2000.
XX 31-MAR-2000; 2000WO-US008621.
PF 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.
PA Shinkets RA, Leach M;
PI WPI; 2000-602362/57.
DR P-PSDB; AAB42993.
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX Claim 5; Page 4692-4693; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; cardiant;
CC anticovulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antihydroid; and antinaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX	Sequence	837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;
Alignment Scores:		
Alignment No.:	1.24e-103	Length: 837
Score:	1273.00	Matches: 245
Percent Similarity:	99.60%	Conservative: 1
Best Local Similarity:	99.19%	Mismatches: 1
Query Match:	94.23%	Indels: 0
DB:	3	Gaps: 0
US-10-043-649-2 (1-261) x AAC77202 (1-837)		
QY	15 SerLeuSerSerSerValGlnGlyGlnGlyProValThrMetGluAlaGluArgSerIys	34
Db	4 AGCTTAGATTCCCTCTGTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAGAGACGAAG	63
QY	35 AlaThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArg	54
Db	64 GCCACAGCCCGCGCCCTGGGCGAGTTCCCGGCAAGTGCCCGCGGAGCTGTGCTGAGA	123
QY	55 LeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpTrpThrValLeuSerGlu	74
Db	124 CTCGGGAGCCATTGACCATCGTCTGTGAGATGAGACTGTGTGACGGTGTGTCTGAA	183
QY	75 ValSerGlyArgGluTrpAsnIleProSerValHisValAlaIysValSerHisGlyTrp	94
Db	184 GTCTCAGGCAGAGAGTATATACATCCCAAGCGTCCACGTGGCCAAAGTCTCCATGGGTGG	243
QY	95 LeuTrpGlnGlyLeuSerArgGluIysAlaGlnGluLeuLeuLeuProGlyAsnPro	114
Db	244 CTGTATGAGGGCCTTGAGCAGGAGAAAGCAGAGACCTGCTGTGTACTGGGAACCT	303
QY	115 GlyGlyAlaPheLeuIleArgGlnSerGlnThrArgArgGlySerTrpSerLeuSerVal	134
Db	304 GGAGGGGCTTCTCTCATCCGGGAGAGCCAGACCAAGAGAGGCTCTTACTCTGTCTAGTC	363
QY	135 ArgLeuSerArgProAlaSerTrpAspArgIleArgHisTrpArgIleHisCysLeuAsp	154
Db	364 CGCCTCAGCCGCCCTGCATCTTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAC	423
QY	155 AsnGlyTrpLeuTrpIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp	174
Db	424 AATGGCTGGCTGACATCTACCGGCGCTCACTTCCCTCACTCCAGGCCCTGTGTGAC	483
QY	175 HisTrpSerGluLeuAlaAspAspIleCysCysLeuLeuIysGluProCysValLeuGln	194
Db	484 CATTACTCTGAGCTGGCGGATGACATCTGTGCTGCTACTCAAGAGAGCCCTGTGTCTGACG	543
QY	195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro	214
Db	544 AGGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACTGTGACTGTGCTGACAGACACCA	603
QY	215 LeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGlu	234
Db	604 CTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTTCTGAAGCTGCCACAGGGGAGGAG	663
QY	235 SerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTrpIleSerLeuAsnAspGlu	254
Db	664 TCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAG	723
QY	255 AlaValSerLeuAspAspAla 261	
Db	724 GCTGTCTCTTTGGATGATGCC 744	
RESULT 8		
ID	ABQ99374	standard; cDNA; 1413 BP.
AC	ABQ99374;	
DT	25-FEB-2003	(first entry)
XX		

DE Human coding sequence SEQ ID 107.

XX Human; expressed sequence tag; EST; chromosome 20;
KW haematopoietic disorder; central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.

OS Homo sapiens.

PN WO200259260-A2.

PD 01-AUG-2002.

PF 16-NOV-2001; 2001WO-US042950.

PR 17-NOV-2000; 2000US-00714936.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-590824/63.

DR N-PSDB; ABP64788.

XX New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.

PS Claim 1; SEQ ID NO 107; 394pp; English.

XX The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotent or pluripotent state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or
CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. haematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 8.64e-98 Length: 1413
XX Score: 1210.50 Matches: 241
XX Percent Similarity: 91.98% Conservative: 0
XX Best Local Similarity: 91.98% Mismatches: 3
XX Query Match: 89.60% Indels: 18
XX DB: 6 Gaps: 1

US-10-043-649-2 (1-261) x ABQ99374 (1-1413)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 54 ATGGAGAGTCTGCCACGACAGAAAGAAATCTTCCCAAGCCCAAGCTTGATCTCTGTC 113
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 114 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 173
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 174 GGCAGTTTCCCGCAGGTGGCCCGGCGGAGCTGTGAGACTCGGGGAGCCATTGACC 233
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 234 ATGCTCTGTGAGATGAGACTGTGTGACCGGTCTGTGAAGTCTCAGGCAGAGAGTAT 293
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 294 AACATCCCCAGCGTCACGTCAGTGGCCAAAGTCTCCCAAGGTGGCTGTATGAGGGCTGAGC 353
QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 354 AGGAGAAAGCAGAGAACTGTGTGTACTCTGGGAACCTTGAGGGGCTTCTCATC 413
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 414 CCGGAGAGCCAGACCCAGAGAGGCTTACTCTCTGACGTCCGCTCAGCCGCTGCA 473
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 474 TCTTGGAACCGGATCAGACACTAAGATCCACTGCTTGACAATGGCTGGCTGTATAC 533
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 534 TCACCGCGCTCACCCTTCTCTCACTCCAGGCCCTGGGGGACCATTTAC----- 581
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200
Db 582 -----TCTGAGGGCTGGCCGCTGCC 602
QY 200 OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220
Db 603 TGGCAAGATATACCCCTACCTGTGACTGTGACAGAGACCACTCAACTGGAAGAGCT 662
QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluG 240
Db 663 GGACAGCTCCCTCCTCTTCTTGAAGCTGCCACAGGGAGAGTCTTCTCAGTGAAGG 722
QY 240 YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAs 260
Db 723 TCTCCGGAGTCCCTCAGCTTCTAATCATCAGCCTGAATGACGAGGCTGTCTTTGATGA 782
QY 260 PAla 261
Db 783 TGCC 786
RESULT 9
AAL44090
ID AAL44090 standard, cDNA; 737 BP.
XX AC AAL44090;
XX DT 03-OCT-2002 (first entry)
XX DE Mouse MARS short isoform protein coding sequence.
XX KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW MARS; tumour suppressor gene; Src-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.
OS Mus sp.

```
XX Key Location/Qualifiers
FH CDS 1..633
FT /*tag= a
FT /product= "Mouse MARS short isoform protein"
XX PN WO200242452-A2.
XX PD 30-MAY-2002.
XX PF 26-NOV-2001; 2001WO-CA001662.
XX PR 27-NOV-2000; 2000CA-02324663.
XX PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX PI Mcglade JC, Loreto MP;
XX DR WPI; 2002-566564/60.
XX DR P-PSDB; AAO15458.
XX
XX PT New isolated modulator of antigen receptor signaling protein or its
PT fragment, useful for treating malignant disorders such as myeloid
PT malignancies, autoimmune disorders and myeloproliferative disorders.
XX
XX PS Claim 9; Page 77; 110pp; English.
XX
XX CC The invention comprises the amino acid and coding sequences of modulator
XX of antigen receptor signaling (MARS) proteins. The MARS protein is a
XX putative tumour suppressor gene and exhibits structural and sequence
XX similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
XX protein sequences of the invention are useful for the treatment of
XX myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
XX disorders, immunosuppression, myeloproliferative disorders and
XX malignancies related to the de-regulation of tyrosine kinases (e.g.
XX breast cancer). The present cDNA sequence encodes a mouse MARS protein
XX
XX SQ Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.87e-97 Length: 737
Score: 1200.50 Matches: 241
Percent Similarity: 92.40% Conservative: 2
Best Local Similarity: 91.63% Mismatches: 1
Query Match: 88.86% Indels: 19
DB: 6 Gaps: 1

US-10-043-649-2 (1-261) x AAL44090 (1-737)
QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCCAAGAGAATAATCTCTGCCCAAGCTTGAGTTCTCTCTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAla 40
Db 61 CAAAGCCAGGGACCTGTGACCATGGAAGCAGAGAACCAAGCCACAGCCGTGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAAGTTCCCGGAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCGTCTCTGAGGATGAGACTGTGTGACGCTGTCTGAAGTCTCAGCAGCAGAGTAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 301 AGGAGAAGACAGAGAACTGCTGTGTACTCTGGAACTCTGAGAGGGCCCTTCTCATC 360
```

```
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGACCAGAGAGGCTCTTACTCTGTCTAGTCCGCCCTCAGCCCGCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 421 TCCTGGACCGGATCAGACACTACAGAGATCCACTGCTGACAAATGGCTGTGTACATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCCCTCAGCTTCCCTCCTCAGCTCCAGGCCCTGTGGACCATTAAC----- 528
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200
Db 529 -----TCTGAGGGCTGGCCGCTCCC 549
QY 200 oglyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220
Db 550 TGGCAAGATATATACCCCTAACCCTGTGACTGTGGGAGGACACCACTCACTGGAAGAGCT 609
QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGl 240
Db 610 GCAAGACTCCCTCTCTTTCTTCTGAAGCTGCCAGGGGAGAGTCTTCTCAGTGAGGG 669
QY 240 yLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA 260
Db 670 TCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGAGCGAGGCTGTCTTTGGATG 729
QY 260 sPALa 261
Db 730 ATGCC 734

RESULT 10
AAS74750
ID AAS74750 standard; cDNA; 2049 BP.
XX
XX AC AAS74750;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #10554.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG10563.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX PS Claim 1; SEQ ID NO 10554; 103bp; English.
XX
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
```

CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 1.18e-90 Length: 2049
Score: 1132.50 Matches: 239
Percent Similarity: 89.89% Conservative: 1
Best Local Similarity: 89.51% Mismatches: 4
Query Match: 83.83% Indels: 23
DB: 5 Gaps: 1

US-10-043-649-2 (1-261) x AAS74750 (1-2049)

QY 1 MetGlySerLeuProSerArgArgIysSerLeuProSerProSerLeuSerSerVal 20
Db 965 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 1024
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40
Db 1025 CAAGGCCAGGAGCCTGTGACCACTGAAAGAGAGAGAAGCAAGCCAGCCGTGCGCTG 1084
QY 41 GlySerPheProAlaGlyIyProAlaGluLeuSerLeuArgLeuGlyIuProLeuThr 60
Db 1085 GGCAGTTTCCCGCAGAGTGCGCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 1144
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 1145 ATCGTCTCTGAGGATGAGACTGGTGAACGGTGTGTCTGAAGTCTCAAGCAGAGAGTAT 1204
QY 81 AsnIleProSerValHisValAlaIysValSerHisGlyTyrPleuTyrGluGlyLeuSer 100
Db 1205 AACATCCCCCAGCGCTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCCTGAGC 1264
QY 101 ArgGluIysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 1265 AGGGAGAAAGCAGAGAACTGCTGTTGTTAAGTGGAAACCTGAGAGGGCCTTCTCATC 1324
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 1325 CGGGAGAGCCAGACCAAGAGAGGCTCTTACTCTCTGTCAATCGCCTCAACCGCCCTGCA 1384
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPleuTyrIle 160
Db 1385 TTCTGGGACCGGATCAGACACTACAGGATTCAGCTTGACATGGCTGGCTGTACATC 1444
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 1445 TCACCCGCGCTCACTTCCCTCACTTCAAGGCCCTGTGTGACCATTAAC----- 1492
QY 181 AspAspIleCysCysLeuLeuGluProCysValLeuGln-ArgAlaGlyProLeuPr 200
Db 1493 -----TCTGAGGGGCTGGCCCGCTCC 1513
QY 200 OGlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpIysGluIle 220

Db 1514 TGGCAAGGATATACCCCTACTGTGACTGTGAGAGAGACCACTCAACTGGAAGAGCT 1573
QY 220 uaspSerSerLeuLeuPheSerGluAlaIaThr-GlyGluGluSerLeu-Leu-SerGlu 239
Db 1574 GGACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGAGAGTCTTCTCAGAGGAG 1633
QY 240 GlyLeuArgGluSer-LeuSerPheTyrIleSer-LeuAsnAspGluAlaValSer-Ileu 258
Db 1634 GGGCTCCGGAGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTTG 1693

RESULT 11

AAL44087
ID AAL44087 standard; cDNA; 1348 BP.
XX
AC AAL44087;
XX
DT 03-OCT-2002 (first entry)
XX
DE Mouse modulator of antigen receptor signalling protein coding sequence.
XX
KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.
XX
OS Mus sp.

XX
FH Key Location/Qualifiers
FT CDS 282..1061
FT /*tag= a
FT /product= "Mouse MARS protein"

XX WO200242452-A2.

XX 30-MAY-2002.

XX 26-NOV-2001; 2001WO-CA001662.

XX 27-NOV-2000; 2000CA-02324663.

XX (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX Mcglade JC, Loreto MP;

XX MPI; 2002-566564/60.

XX P-PSDB; AAO15456.

XX
PT New isolated modulator of antigen receptor signalling protein or its
PT fragment, useful for treating malignant disorders such as myeloid
PT malignancies, autoimmune disorders and myeloproliferative disorders.
XX

XX Claim 10; Fig 1A; 110bp; English.

XX
XX The invention comprises the amino acid and coding sequences of modulator
CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present cDNA sequence encodes a mouse MARS protein
XX

SQ Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.73e-82 Length: 1348
Score: 1032.00 Matches: 209
Percent Similarity: 85.88% Conservative: 16

Best Local Similarity: 79.77% Mismatches: 33
Query Match: 76.39% Indels: 4
DB: 6 Gaps: 3

US-10-043-649-2 (1-261) x AAL44087 (1-1348)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 282 ATGGGAAGTTGTTCACAGAGAGGAAACC--TCCAGCCCCAGCCCCAGCTCTGTGT 338
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 339 CCAGACCAGGAACCGTGTCTCATGCAACAGAAAGACACAAGGTACAGCTGTGCCCCCTG 398
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 399 GGCAAGTTTCCAGCAGGTCAACAGCCAGACTATCTGTGAGACTCGGGAGCCGCTGACC 458
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 459 ATCATCTCTGAGGATGAGATGTGTGACAGCTCCAGTCGGAAGTCTCAGGCGAGAGTAC 518
QY 81 AsnIleProSerValHisValAlaAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 519 CACATGCCCAAGTGTGTATGTGGCTAAAGTCGCCCAAGGGGTGGCTGTACGAGGGCTGAGC 578
QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyAlaAlaPheLeuIle 120
Db 579 CGGGAGAAAGCCGAGGAACTACTCTGTGTACCTGGGAACCCGAGGGGCTTCTCTCATC 638
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 639 CGGGAGAGCCAGACCAAGAGAGGCTGTATCTCCCTGCTCCGACTCAGCCGCTGCA 698
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 699 TCTTGGACCGGATCAGACACTACAGATACAGCCGCTTGACATGCGCTGTACATC 758
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 759 TCACCTGCCTCACCCTTCCCTCCTCACTCCACGCGCTTGCTGAGCATTACTCTGAGCTAGCA 818
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 819 GATGGCATCTGCTGTCTCCCTCAGGGAGCCGTGTCTCTGAGAAAGCTTGGGCACTACCT 878
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 879 GGCAAAGATACACCTCACTGTGACTGTGCAACATCATCACTAAATTGGAAGAAAGCTG 938
QY 221 AspSerSerLeuLeuPheSerGluAla--AlaThrGlyGluGluSerLeuSerGlu 239
Db 939 GACCGCAGCCTCCTGTTCTGGAAGCACCTCGAGTGGGGAGGCATCTCTGCTCAGTGAG 998
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 999 GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCTGGCTGAGGAC-----CCCTTGAT 1052
QY 260 AspAla 261
Db 1053 GATGCT 1058

RESULT 12
ABQ98670 standard; DNA; 763 BP.
XX ABQ98670;
AC
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF477 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;

KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PE 30-MAY-2001; 2001US-00867550.
XX
PR 30-MAY-2000; 2000US-0208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
XX (LAWD/) LAW D.
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX WPI; 2002-626554/67.
DR P-PSDB; ABP64107.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
PS Claim 2; SEQ ID NO 953; 78bp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=2002082206
XX
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.36e-64 Length: 763
Score: 830.00 Matches: 159
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.44% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x ABQ98670 (1-763)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 286 ATGGGAAGTTGTCCACAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 345
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 346 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCAAGGCCACAGCCGTGCTG 405
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 406 GGCAAGTTTCCGCGAGGTGGCCCGGCGAGCTGTCTGAGACTCGGGAGCCATTGACC 465
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80

PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009126.
XX
PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 21302; 106bp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 864 BP; 191 A; 259 C; 242 G; 165 T; 0 U; 7 Other;
XX
Alignment Scores:
Pred. No.: 4 9e-48 Length: 864
Score: 647.00 Matches: 127
Percent Similarity: 99.22% Conservative: 1
Best Local Similarity: 98.45% Mismatches: 1
Query Match: 47.89% Indels: 0
DB: 5 Gaps: 0
US-10-043-649-2 (1-261) x ADL63090 (1-864)
QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerSerVal 20
Db 450 ATGGGAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTGAGTCTCTCTGTC 509
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40

Db 510 CAAGCCAGGAGCTGTGACCATGGAAAGCAGAGAAAGCCACAGCCGTGCCCTG 569
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 570 GGCAGTTTCCCGGAGGTGGCCCGCCGAGCTGTGAGACTCGGAGGCAATTGACC 629
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 630 ATCGTCTCTGAGAGTGAAGACTGTGTGACCGGTGTCTGAAGTCTCAGGCAGAGATAT 689
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 690 AACATCCCCAGCGTCCACGTGGCCCAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGC 749
QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 750 AGGAGAAAGCAGAGGAACTGCTGTTTACTGTGGAAACCTGAGGGGCTTCTCATC 809
QY 121 ArgGluSerGlnThrArgArgGlySer 129
Db 810 CGGAGAGCCAGACCAAGAAAGAGTCC 836
RESULT 15
ABQ99151
ID ABQ99151 standard; DNA; 875 BP.
XX
AC ABQ99151;
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF958 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-00867550.
XX
PR 30-MAY-2000; 2000US-0208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPE/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI; 2002-626554/67.
XX
P-PSDB; ABP64588.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
PS Claim 2; SEQ ID NO 1915; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological

CC condition associated with an ORFX-associated disorder, e.g. cancer, blood
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=20020082206

XX
SQ Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.3e-42	Length:	875
Score:	586.00	Matches:	112
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	43.38%	Indels:	0
DB:	6	Gaps:	0

US-10-043-649-2 (1-261) x ABQ99151 (1-875)

QY	150	ILEHISCYSLLEUASPANGLYTRPLEUTYRILLESERPROARGLEUTHRPHEPROSERLEU	169
DB	4	ATCCACTGCCCTTGACATGGCTGGCTGTACATCTCAACCGCCTCACCTTCCCTCACTC	63
QY	170	GLNALALEUVALASPHISTYRSERGLULEUALASPAPILECYSLEULEULYSGLU	189
DB	64	CAGGCCCTGGTGGACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAAGAG	123
QY	190	PROCYSVALLLEUGLARGALAGLYPROLEUPROGLYLYSASPILPROLEUPROVALTHR	209
DB	124	CCCTGTGTCTGCAGAGGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACTGTGACT	183
QY	210	VALGLNARGTHRPROLEUASNTRPLYSGLULEUASPSESRLEULEUPHESERGIUALA	229
DB	184	GTGCAGAGGACACCACTCACTGAAAGAGCTGACAGCTCCCTGTTTCTGAAGCT	243
QY	230	ALATHRGLYGLUGLUSERLEUENSERGLUGLYLEUARGGLUSERLEUSERPHEITYRILE	249
DB	244	GCCACAGGGAGAGGTCTTCTCACTGAGGGTCTCCGGGAGTCCCTCAGCTTCTTACATC	303
QY	250	SERLEUASNAPGLUALAVASERLEUASPAPALA	261
DB	304	AGCCTGAATGACGAGGCTGTCTTGTGATGATGCC	339

Search completed: November 17, 2004, 00:59:18
Job time : 539 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2004, 11:56:01 ; Search time 509 Seconds

(without alignments)
8106.179 Million cell updates/sec

Title: US-10-043-649-1

Perfect score: 786

Sequence: 1 atgggaagctctgccacgacg.....tctcttgatgatgcctag 786

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	786	100.0	786	6	AAL44089	Aa144089 Human mod
2	786	100.0	786	6	ABQ74343	Abq74343 Human Src
3	786	100.0	2567	6	AAD43980	Aad43980 Human Src
4	786	100.0	3646	10	ADF90741	Adf90741 Human hep
5	785	99.9	2788	10	ADF82458	Adf82458 Leukaemia
6	784.4	99.8	1183	6	ABK61465	Abk61465 Human cDN
7	741.8	94.4	837	3	AAC77202	Aac77202 Human ORF
8	672.8	85.6	1413	6	ABQ99374	Abq99374 Human cod
9	663.4	84.4	737	6	AAL44090	Aal44090 Mouse MAR
10	609.8	77.6	2049	5	AAS74750	Aas74750 DNA encod
11	537.8	68.4	1348	6	AAL44087	Aal44087 Mouse mod
12	478	60.8	763	6	ABQ98670	Abq98670 Human ORF
13	405	51.5	603	5	AAS74748	Aas74748 DNA encod
14	381	48.5	864	5	ADL63090	Adl63090 Human ova
15	341	43.4	875	6	ABQ99151	Abq99151 Human ORF
16	215	27.4	320	5	ADL43424	Adl43424 Human ova
17	192.8	24.5	211	5	AAS70181	Aas70181 DNA encod
18	192	24.4	444	6	ABQ98669	Abq98669 Human ORF
19	160.2	20.4	2109	4	AAS02049	Aas02049 DNA encod
20	159	20.2	2665	6	ABL65189	Ab165189 Lung canc
21	159	20.2	2665	6	ABK83738	Abk83738 Human cDN

22	159	20.2	2665	8	ACC81091	Acc81091 Human Src
23	159	20.2	2665	12	ADL83115	Adl83115 Human PRO
24	159	20.2	2665	12	ADP12833	Adp12833 Reference
25	159	20.2	3452	12	ADQ22520	Adq22520 Human sof
26	159	20.2	3756	8	ABX62975	Abx62975 Human act
27	141.8	18.0	432	4	AAL12879	Aal12879 Probe #28
28	141.8	18.0	432	4	ABA54580	Abas4580 Human foe
29	141.8	18.0	432	4	AAL134236	Aal134236 Probe #29
30	141.8	18.0	432	4	ABA44128	Abas44128 Human bre
31	141.8	18.0	432	4	ABA24363	Abas24363 Probe #28
32	141.8	18.0	432	4	AAK28314	Aak28314 Human bon
33	141.8	18.0	432	4	AAK02872	Aak02872 Human bra
34	141.8	18.0	432	4	ABS27912	Abs27912 Human liv
35	141.8	18.0	432	5	AAL02797	Aal02797 Probe #27
36	141.8	18.0	432	6	ABS02823	Abs02823 Human gen
37	141.8	18.0	448	4	AAL14520	Aal14520 Probe #44
38	141.8	18.0	448	6	ABS04499	Abs04499 Human gen
39	133.2	16.9	152	5	ADL71889	Adl71889 Human ova
40	133.2	16.9	152	5	ADL37040	Adl37040 Human ova
41	124.6	15.9	1924	12	ADL04091	Adl04091 Human HCK
42	124.6	15.9	1926	6	ABK83940	Abk83940 Human cDN
43	124.6	15.9	1926	12	ADJ71658	Adj71658 Human NOV
44	124.6	15.9	2015	6	ABL66673	Ab166673 Lung canc
45	124.6	15.9	2015	6	ABK83939	Abk83939 Human cDN

ALIGNMENTS

RESULT 1	
AAL44089	
ID	AAL44089 standard; cDNA, 786 BP.
AC	AAL44089;
DT	03-OCT-2002 (first entry)
DE	Human modulator of antigen receptor signalling protein coding sequence.
XX	
KW	Human; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW	MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW	myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW	immunosuppression; myeloproliferative disorder; breast cancer.
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..786
FT	/*tag= a
FT	/product= "Human MARS protein"
XX	
PN	WO200242452-A2.
XX	
PD	30-MAY-2002.
XX	
PF	26-NOV-2001; 2001WO-CA001662.
XX	
PR	27-NOV-2000; 2000CA-02324663.
XX	
PA	(HOSP-) HOSPITAL FOR SICK CHILDREN.
XX	
PI	McGlade JC, Loreto MP;
XX	
DR	WPI: 2002-566564/60.
DR	P-PSDB; AAO15457.
XX	
PT	New isolated modulator of antigen receptor signaling protein or its
PT	fragment, useful for treating malignant disorders such as myeloid
PT	malignancies, autoimmune disorders and myeloproliferative disorders.
XX	
PS	Claim 12; Page 75; 110pp; English.
XX	
CC	The invention comprises the amino acid and coding sequences of modulator

CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present cDNA sequence encodes a human MARS protein
XX
SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Query Match 100.0%; Score 786; DB 6; Length 786;
Best Local Similarity 100.0%; Pred. No. 1.1e-197;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCCCAGCAGAAATAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
Db 1 ATGGGAAGTCTGCCAGCAGAAATAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 120
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 120
QY 121 GGCAAGTTTCCCGCAGGTGGCCCCGGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
Db 121 GGCAAGTTTCCCGCAGGTGGCCCCGGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
QY 181 ATGCTCTGTAGGATGAGACTGTGGACGGTGTGCTGTAAGTCTCAGGCAGAGAGATAT 240
Db 181 ATGCTCTGTAGGATGAGACTGTGGACGGTGTGCTGTAAGTCTCAGGCAGAGAGATAT 240
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCAAGGGTGGCTGTATGAGGGCCCTGAGC 300
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCAAGGGTGGCTGTATGAGGGCCCTGAGC 300
QY 301 AGGGAAGAACAGAGAACTGTTGTTACCTGGAAACCTGGAGGGCCCTTCTCATC 360
Db 301 AGGGAAGAACAGAGAACTGTTGTTACCTGGAAACCTGGAGGGCCCTTCTCATC 360
QY 361 CGGAGAGCCAGACCCAGAGAGAGGCTTACTCTCTGTCAATCCGCTCAGCCGCCCTGCA 420
Db 361 CGGAGAGCCAGACCCAGAGAGAGGCTTACTCTCTGTCAATCCGCTCAGCCGCCCTGCA 420
QY 421 TCCTGGGACCCGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACATC 480
Db 421 TCCTGGGACCCGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACATC 480
QY 481 TCACCGGCGCTCACCCTTCCCTCACTCCAGGCGCTGTGACCATTAATCTGAGCTGGCG 540
Db 481 TCACCGGCGCTCACCCTTCCCTCACTCCAGGCGCTGTGACCATTAATCTGAGCTGGCG 540
QY 541 GATGACATCTGCTGCTTCACTCAAGAGCGCTGTGCTGCTGCAAGGGCTGGCCGCTCCCT 600
Db 541 GATGACATCTGCTGCTTCACTCAAGAGCGCTGTGCTGCTGCAAGGGCTGGCCGCTCCCT 600
QY 601 GGCAGAGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGGAAGAGCTG 660
Db 601 GGCAGAGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGGAAGAGCTG 660
QY 661 GACAGTCCCTCTCTGTTTCTGAAGTCCACAGAGGAGAGTCTTCTCAGTAGGGT 720
Db 661 GACAGTCCCTCTCTGTTTCTGAAGTCCACAGAGGAGAGTCTTCTCAGTAGGGT 720
QY 721 CTCCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTTGATGAT 780
Db 721 CTCCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTTGATGAT 780
QY 781 GCCTAG 786
Db 781 GCCTAG 786

RESULT 2

AB074343
ID AB074343 standard; cDNA; 786 BP.
XX
AC AB074343;
XX
DT 15-OCT-2002 (first entry)
XX
DE Human Src-like inhibitory molecule (SLIM) encoding cDNA.
XX
KW Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;
KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;
KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;
KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;
KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..786
FT /*tag= a
FT /product= "SLIM"
FT /note= "Src-like inhibitory molecule"
XX
PD WO200255707-A2.
XX
PD 18-JUL-2002.
XX
PF 10-JAN-2002; 2002WO-US000718.
XX
PR 10-JAN-2001; 2001US-0260953P.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;
PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;
XX
DR MPI; 2002-575432/61.
DR P-PSDB; ABP52187.
XX
PT New src-like inhibitory molecule protein, useful for treating
PT immunodeficiency disorders and inflammatory disorders, comprises N-
PT terminal myristylation sequence, SH2 domain and/or SH3 domain.
XX
PS Claim 3; Fig 2A; 91pp; English.
XX
CC The present sequence encodes the human Src-like inhibitory molecule
CC (SLIM) protein (I). The present invention describes a SLIM protein
CC comprising an N-terminal myristylation sequence, an N-terminal SH2
CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising
CC an N-terminal myristylation sequence and an N-terminal SH2 domain which
CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and
CC anti-HIV activities, and can be used as a modulator of lymphocyte
CC activation, and of ubiquitination of a Cbl target protein, and in gene
CC therapy. (I) is useful for screening a bioactive agent capable of binding
CC to SLIM. (I) is also useful for screening a bioactive agent capable of
CC modulating SLIM binding. (I) or its fragments is useful in the study or
CC in the treatment of conditions which involves this function or
CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or
CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it
CC (II) is useful for modulating leukocyte and/or platelet activation, for
CC modulating antigen receptor-induced signalling and activation in
CC leukocyte and/or platelets and for modulating antigen receptor-induced
CC signalling and activation in lymphocytes and/or mast cells. (I) or (II)
CC is also useful for modulating the basal activity of lymphocytes. (I) or
CC (II) is useful in the treatment of immunodeficiency disorders, such as
CC acquired immunodeficiency syndrome (AIDS), for the prevention and
CC treatment of acute inflammatory disorders, chronic inflammatory
CC disorders, autoimmune disorder and transplant rejection
XX
SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Query Match 100.0%; Score 786; DB 6; Length 786;

Best Local Similarity 100.0%; Pred. No. 1.1e-197;		Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	60
Db	1	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	60
QY	61	CAAGGCCAGGGACCTGTGACCATGGAAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTG	120
Db	61	CAAGGCCAGGGACCTGTGACCATGGAAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTG	120
QY	121	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC	180
Db	121	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC	180
QY	181	ATCGTCTCTGAGATGAGACTGTGGACCGTGTGTAAGTCTCAGGCGAGAGATAT	240
Db	181	ATCGTCTCTGAGATGAGACTGTGGACCGTGTGTAAGTCTCAGGCGAGAGATAT	240
QY	241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATAGGGCCTGAGC	300
Db	241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATAGGGCCTGAGC	300
QY	301	AGGGAGAAAGCAGAGAACTGTGTGTACCTGGAAACCTGGAGGGGCTTCCTCATC	360
Db	301	AGGGAGAAAGCAGAGAACTGTGTGTACCTGGAAACCTGGAGGGGCTTCCTCATC	360
QY	361	CGGGAGAGCCAGACGAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCTTCA	420
Db	361	CGGGAGAGCCAGACGAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCTTCA	420
QY	421	TCCTGGACCGGATGACACTACAGAGATCCACTGCTTGACATGGCTGGTGTACATC	480
Db	421	TCCTGGACCGGATGACACTACAGAGATCCACTGCTTGACATGGCTGGTGTACATC	480
QY	481	TCACCGCGCTTACCTTCCCTCCTACTCCAGGCGCTGTGGACCATTAATTAGCTGGCG	540
Db	481	TCACCGCGCTTACCTTCCCTCCTACTCCAGGCGCTGTGGACCATTAATTAGCTGGCG	540
QY	541	GATGACATCTGCTGCTTCTCAAGAGAGCCCTGTGTCTGAGAGGGCTGGCCGCTCCCT	600
Db	541	GATGACATCTGCTGCTTCTCAAGAGAGCCCTGTGTCTGAGAGGGCTGGCCGCTCCCT	600
QY	601	GGCAAGGATATACCCCTTACTGTGACTGTGACAGAGACCACTCAACTGGAAGAAGCTG	660
Db	601	GGCAAGGATATACCCCTTACTGTGACTGTGACAGAGACCACTCAACTGGAAGAAGCTG	660
QY	661	GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCAGTAGGGT	720
Db	661	GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCAGTAGGGT	720
QY	721	CTCCGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT	780
Db	721	CTCCGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT	780
QY	781	GCCTAG 786	
Db	781	GCCTAG 786	

RESULT 3

AAD43980	standard; cDNA; 2567 BP.
ID	AAD43980
XX	
AC	AAD43980;
XX	
DT	13-DEC-2002 (first entry)
XX	
DE	Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.
XX	
KW	Human; SH2/SH3-domain-containing adapter; Src-like Adapter Protein-2;
KW	SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;
KW	neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;

KW	psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;
KW	Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;
KW	multiple sclerosis; asthma; acute respiratory distress syndrome;
KW	pulmonary disorder; dermatological; neuroprotective; gene; ss.
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	415..1200
FT	/tag= a
FT	/product= "Human SLAP-2"
XX	
PN	WO200242457-A1.
XX	
PD	30-MAY-2002.
XX	
PF	20-NOV-2001; 2001WO-US043367.
XX	
PR	22-NOV-2000; 2000US-0252545P.
XX	
PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX	
PI	Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;
XX	
DR	WPI; 2002-463632/49.
DR	P-PSDB; AAE26357.
XX	
PT	Novel substantially purified human SH2/SH3-domain-containing adapter
PT	polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic
PT	intervention in immunological and inflammatory disorders and cancer.
XX	
PS	Claim 2; Fig 1; 85pp; English.
XX	
CC	The invention relates to a substantially purified human SH2/SH3-domain-
CC	containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-
CC	2). The invention is useful for treating an immune disorder involving
CC	hyperactivity of B- or T- lymphocytes in a mammal. The invention is
CC	useful for screening for antagonists or inhibitors of the interaction of
CC	hSLAP-2 with cellular signalling compounds, for diagnosing, treating or
CC	preventing diseases or disorders associated with aberrant or uncontrolled
CC	cellular signal transduction, for determining those cellular signalling
CC	molecules which associate with hSLAP-2 and which provide critical signals
CC	for cell activation, and as effectors in methods to affect T- cell
CC	activation. The invention is useful in screening assays to identify and
CC	detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for
CC	potential use to treat autoimmune diseases which may be caused by
CC	hyperactivated B cells, as well as to treat diseases which may be caused
CC	by hyperactivated T cells, in addition to other immune system related
CC	conditions, diseases, or disorders, T-cell and B-cell neoplasms, rheumatoid arthritis,
CC	inflammation disorders, diseases and conditions, osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's
CC	and ulcerative colitis), allergies, particularly those involving
CC	hyperactivity of B-cells and T- cells, or other immune cells, such as
CC	mast cells or eosinophils, autoimmune diseases such as systemic lupus
CC	erythematosus and multiple sclerosis, pulmonary diseases including
CC	asthma, acute respiratory distress syndrome, and chronic obstructive
CC	pulmonary disorder, tissue/ organ rejection and cancer. The invention is
CC	useful in gene therapy. The present sequence is human SLAP-2 cDNA
XX	
SQ	Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Query Match	100.0%; Score 786; DB 6; Length 2567;
Best Local Similarity	100.0%; Pred. No. 1.5e-197;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

QY	1	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	60
Db	415	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	474
QY	61	CAAGGCCAGGGACCTGTGACCATGGAAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTG	120
Db	475	CAAGGCCAGGGACCTGTGACCATGGAAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTG	534

QY 121 GGCAGTTTCCCGGAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
|||
Db 535 GGCAGTTTCCCGGAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 594
QY 181 ATCGTCTGTAGGATGAGACTGTGTGACGGTGTGCTGAAGTCTCAGGCAAGAGTAT 240
|||
Db 595 ATCGTCTGTAGGATGAGACTGTGTGACGGTGTGCTGAAGTCTCAGGCAAGAGTAT 654
QY 241 AACATCCCCAGCGTCCAGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 300
|||
Db 655 AACATCCCCAGCGTCCAGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 714
QY 301 AGGAGAAAGCAGAGGAAGTGTGTGTACCTGGGAACCTTGAGGGGCTTCTCATC 360
|||
Db 715 AGGAGAAAGCAGAGGAAGTGTGTGTACCTGGGAACCTTGAGGGGCTTCTCATC 774
QY 361 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCCGCTGCA 420
|||
Db 775 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCCGCTGCA 834
QY 421 TCCTGGAGCCGATCAGACTACAGAGATCCACTGCTTGACAATGGCTGGCTGTACATC 480
|||
Db 835 TCCTGGAGCCGATCAGACTACAGAGATCCACTGCTTGACAATGGCTGGCTGTACATC 894
QY 481 TCACCGCGCTCACCCTCCCTCACTCCAGGCCCTGGTGAACCATTACTCTGAGCTGGC 540
|||
Db 895 TCACCGCGCTCACCCTCCCTCACTCCAGGCCCTGGTGAACCATTACTCTGAGCTGGC 954
QY 541 GATGACATCTGTGCTCTCAAGAGAGCCCTGTGTCTGAGAGGGCTGGCCGCTCCCT 600
|||
Db 955 GATGACATCTGTGCTCTCAAGAGAGCCCTGTGTCTGAGAGGGCTGGCCGCTCCCT 1014
QY 601 GGCAGGATATACCCCTACTCTGTGACTGTGACAGAGACACCACTCAACTGGAAAAGACTG 660
|||
Db 1015 GGCAGGATATACCCCTACTCTGTGACTGTGACAGAGAGACACCACTCAACTGGAAAAGACTG 1074
QY 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTTCTCAGTGAAGGT 720
|||
Db 1075 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTTCTCAGTGAAGGT 1134
QY 721 CTCGGGAGTCCCTCAGCTTCTCATCAAGCTGAATGACGAGGCTGTCTCTTGATGAT 780
|||
Db 1135 CTCGGGAGTCCCTCAGCTTCTCATCAAGCTGAATGACGAGGCTGTCTCTTGATGAT 1194
QY 781 GCCTAG 786
|||
Db 1195 GCCTAG 1200

RESULT 4
ADF90741/C
ID ADF90741 standard; DNA, 3646 BP.
XX
AC ADF90741;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human hepatic-fibrosis disease marker SEQ ID 203.
XX
KW Hepatic fibrosis; marker; chronic hepatitis; liver cirrhosis;
KW hepatic carcinoma; human; ds.
XX
OS Homo sapiens.
XX
PN JF2003259877-A.
XX
PD 16-SEP-2003.
XX
PF 11-MAR-2002; 2002JP-00065013.
XX
PR 11-MAR-2002; 2002JP-00065013.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.

XX
DR WPI; 2003-821598/77.
XX
PT Hepatic fibrosis disease markers comprising polynucleotides or
PT antibodies, useful for improved diagnosis, screening and developing drugs
PT to treat hepatitis, to control cirrhosis and carcinoma.
XX
PS Claim 1; SEQ ID NO 203; 313bp; Japanese.
XX
CC The present invention relates to hepatic-fibrosis disease markers
CC (ADF90539-ADF90871) and related proteins (ADF90872-ADF90917). The
CC sequences are useful for detecting and treating hepatic fibrosis caused
CC by alcohol consumption, virus infection, etc., and the associated chronic
CC hepatitis, etc. leading to liver cirrhosis and hepatic carcinoma. The
CC markers allow the cause of hepatic fibrosis to be clarified (diagnostic
CC precision), so more suitable treatments can be developed and given.
XX
SQ Sequence 3646 BP; 782 A; 954 C; 1031 G; 877 T; 0 U; 2 Other;

Query Match 100.0%; Score 786; DB 10; Length 3646;
Best Local Similarity 100.0%; Pred. No. 1.7e-197;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGTCTGCCAGAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCCTGTG 60
|||
Db 3202 ATGGAGTCTGCCAGAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCCTGTG 3143
QY 61 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCAAGCCAGCCGTGCCCTG 120
|||
Db 3142 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCAAGCCAGCCGTGCCCTG 3083
QY 121 GGCAGTTTCCCGGAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
|||
Db 3082 GGCAGTTTCCCGGAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 3023
QY 181 ATCGTCTGTAGAGTGAAGTGTGACGGTGTGCTGTGAACTCTCAGGCAAGAGTAT 240
|||
Db 3022 ATCGTCTGTAGAGTGAAGTGTGACGGTGTGCTGTGAACTCTCAGGCAAGAGTAT 2963
QY 241 AACATCCCCAGCGTCCAGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300
|||
Db 2962 AACATCCCCAGCGTCCAGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 2903
QY 301 AGGAGAAAGCAGAGAACTGTGTGTACCTGGGAACCTTGAAGGGCTTCTCATC 360
|||
Db 2902 AGGAGAAAGCAGAGAACTGTGTGTACCTGGGAACCTTGAAGGGCTTCTCATC 2843
QY 361 CGGAGAGCCAGACAGAGAGGCTTTACTCTGTGACGTCCGCTCAGCCGCTGCA 420
|||
Db 2842 CGGAGAGCCAGACAGAGAGGCTTTACTCTGTGACGTCCGCTCAGCCGCTGCA 2783
QY 421 TCCTGGAGCCGATCAGACTACAGAGATCCACTGCTTGACAATGGCTGCTATCATC 480
|||
Db 2782 TCCTGGAGCCGATCAGACTACAGAGATCCACTGCTTGACAATGGCTGCTATCATC 2723
QY 481 TCACCGGCTCACTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTTGAGTGGC 540
|||
Db 2722 TCACCGGCTCACTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTTGAGTGGC 2663
QY 541 GATGACATCTGCTGCTTCAAGAGCCCTGTGTCTGACAGGGCTGGCCGCTCCCT 600
|||
Db 2662 GATGACATCTGCTGCTTCAAGAGCCCTGTGTCTGACAGGGCTGGCCGCTCCCT 2603
QY 601 GGCAGGATATACCCCTACTGTGACTGTGACAGAGACACCACTCACTGGAAGAGCTG 660
|||
Db 2602 GGCAGGATATACCCCTACTGTGACTGTGACAGAGACACCACTCACTGGAAGAGCTG 2543
QY 661 GACAGCTCCCTCTGTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT 720
|||
Db 2542 GACAGCTCCCTCTGTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT 2483
QY 721 CTCGGGAGTCCCTCAGCTTCTCATCAAGCTGAATGACGAGGCTGTCTCTTGATGAT 780
|||

Db 2482 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGAGGCTGTCTCTTGATGAT 2423
QY 781 GCCTAG 786
Db 2422 GCCTAG 2417
RESULT 5
ADF82458
ID ADF82458 standard; DNA; 2788 BP.
XX ADF82458;
AC ADF82458;
XX 26-FEB-2004 (first entry)
DT 26-FEB-2004 (first entry)
XX Leukaemia-related DNA sequence #3014.
DE Leukaemia-related DNA sequence #3014.
XX Cytostatic; Gene therapy; leukaemia; ss.
KM Cytostatic; Gene therapy; leukaemia; ss.
XX Unidentified.
OS Unidentified.
XX WO2003039443-A2.
PN 15-MAY-2003.
XX 15-MAY-2003.
PD 15-MAY-2003.
XX 04-NOV-2002; 2002WO-EP012303.
PF 04-NOV-2002; 2002WO-EP012303.
XX 05-NOV-2001; 2001EP-00126244.
PR 30-APR-2002; 2002EP-00009758.
XX 30-APR-2002; 2002EP-00009758.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (UYLU-) UNITV LUDWIG MAXIMILIANS.
PA (HAFE/) HAFERLACH T.
PA (SCHO/) SCHOCH C.
PA (KERN/) KERN W.
XX (KERN/) KERN W.
PI Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;
PI Eils R, Brors B, Mergenthaler S;
XX WPI; 2003-505037/47.
DR WPI; 2003-505037/47.
XX
XX Determining the subtype of leukemia cells and whether a patient sample
PT contains leukemia cells or other cells, useful for treating leukemia,
PT comprises determining the expression profile of a group of markers in a
PT patient sample.
XX
XX Disclosure; SEQ ID NO 3014; 2938bp; English.
PS Disclosure; SEQ ID NO 3014; 2938bp; English.
XX
XX The present invention relates to a method (M1) for determining the
CC subtype of leukaemia cells and whether a patient sample contains
CC leukaemia cells. The method comprises determining the expression profile
CC of a group of markers in a patient sample. The method is useful for
CC determining the presence of leukaemia cells, its types or subtypes, and
CC for the preparation of a medicament for treating leukaemia.
XX
XX Sequence 2788 BP; 663 A; 764 C; 700 G; 595 T; 0 U; 66 Other;
SQ Sequence 2788 BP; 663 A; 764 C; 700 G; 595 T; 0 U; 66 Other;
Query Match 99.9%; Score 785; DB 10; Length 2788;
Best Local Similarity 99.9%; Pred. No. 2.8e-197;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGAAGTCTGCCAGAGAAGAAATCTCTGCCAGCCCAAGCTTGAGTTCTCTGTC 60
Db 387 ATGGGAAGTCTGCCAGAGAAGAAATCTCTGCCAGCCCAAGCTTGAGTTCTCTGTC 446
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGSCCTG 120
Db 447 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGSCCTG 506
QY 121 GGCAAGTTCCCGCAGGTGCCCCGGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
Db 507 GGCAAGTTCCCGCAGGTGCCCCGGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 566

QY 181 ATGCTCTTGAGGATGAGACTGTGTGAGCGGTGCTGTCTGAAGTCTCAGGCAGAGATAT 240
Db 567 ATGCTCTTGAGGATGAGACTGTGTGAGCGGTGCTGTCTGAAGTCTCAGGCAGAGATAT 626
QY 241 AACATCCCCAGCGTCCAGCTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
Db 627 AACATCCCCAGCGTCCAGCTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 686
QY 301 AGGAGAAAGCAGAGAACTGCTGTGTGTACCTGGGAACCTGGAGGGCGCTTCTCATC 360
Db 687 AGGAGAAAGCAGAGAACTGCTGTGTGTACCTGGGAACCTGGAGGGCGCTTCTCATC 746
QY 361 CGGAGAGCCAGACCAAGAGAGGCTTCTACTCTCTGTCAGTCCGCCCTCAGCCCGCTGCA 420
Db 747 CGGAGAGCCAGACCAAGAGAGGCTTCTACTCTCTGTCAGTCCGCCCTCAGCCCGCTGCA 806
QY 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGTGATATC 480
Db 807 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGTGATATC 866
QY 481 TCACCGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGGAACCATTAATCTAGAGTGGCG 540
Db 867 TCACCGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGGAACCATTAATCTAGAGTGGCG 926
QY 541 GATGACATCTGCTGCTCACTCAAGAGCGCTGTGCTGCAAGAGGGCTGGCCGCTCCCT 600
Db 927 GATGACATCTGCTGCTCACTCAAGAGCGCTGTGCTGCAAGAGGGCTGGCCGCTCCCT 986
QY 601 GGCAAGATATACCCCTAAGCTGTGACTGTGCAAGAGCAACCACTCAACTGGAAGAAGCTG 660
Db 987 GGCAAGATATACCCCTAAGCTGTGACTGTGCAAGAGCAACCACTCAACTGGAAGAAGCTG 1046
QY 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT 720
Db 1047 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT 1106
QY 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGATGAT 780
Db 1107 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGATGAT 1166
QY 781 GCCTAG 786
Db 1167 GCCTAG 1172
RESULT 6
ABK61465
ID ABK61465 standard; cDNA; 1183 BP.
XX ABK61465;
AC ABK61465;
XX 18-JUN-2002 (first entry)
DT 18-JUN-2002 (first entry)
XX Human cDNA encoding protein NOV13.
DE Human cDNA encoding protein NOV13.
XX
XX Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
KW cell signal processing disorder; metabolic pathway modulation disorder;
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;
KW uterus cancer; immune response; graft-versus-host disease;
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW Albright hereditary osteodystrophy.
XX Homo sapiens.
XX
XX WO200216599-A2.
PN WO200216599-A2.
XX
XX 28-FEB-2002.
PD 28-FEB-2002.
XX
XX 27-AUG-2001; 2001WO-US026510.
PF 27-AUG-2001; 2001WO-US026510.
XX
XX 25-AUG-2000; 2000US-0228191P.
PR 25-AUG-2000; 2000US-0228191P.
XX
XX 08-FEB-2001; 2001US-0267300P.
PR 08-FEB-2001; 2001US-0267300P.

PR 20-FEB-2001; 2001US-0269961P.
PR 20-MAR-2001; 2001US-0277337P.
XX
PA (CURA-) CURAGEN CORP.
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA,
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX
DR WPI; 2002-280937/32.
DR P-PSDB; AAU91308.

PT New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
PS Claim 1; Page 98; 263pp; English.

XX
CC The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOVX-14,
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
CC and antibody against it, are useful for treating or preventing (e.g. by
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC atherosclerosis, a disorder related to cell signal processing and
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC and nucleic acids are also useful for determining the presence of
CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC especially useful in therapeutic or prophylactic applications for
CC disorders associated with aberrant NOVX expression or activity, e.g.
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC cancer), immune response, graft-versus-host disease, acquired
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC congenital heart defects, multiple sclerosis, inflammation or Albright
CC hereditary osteodystrophy and many other diseases listed in the
CC specification. The DNA encoding the protein is useful in gene therapy for
CC treating the conditions. This is also useful in detection assays,
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC for developing a powerful assay system for functional analysis of various
CC human disorders, as well as in diagnostic applications. The present
CC sequence encodes a NOVX protein

XX
SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match 99.8%; Score 784.4; DB 6; Length 1183;
Best Local Similarity 99.9%; Pred. No. 3.2e-197;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 60
Db 398 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 457
QY 61 CAAGGCCAGGGACCTGTGACCATGGAAGCAGAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 120
Db 458 CAAGGCCAGGGACCTGTGACCATGGAAGCAGAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 517
QY 121 GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
Db 518 GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 577
QY 181 ATCGTCTGAGGATGAGACTGTGGAACGGTGTCTGAAAGTCTCAGGCAGAGAGTAT 240
Db 578 ATCGTCTGAGGATGAGACTGTGGAACGGTGTCTGAAAGTCTCAGGCAGAGAGTAT 637
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGCGGTGGCTGTATGAGGGCCTGAGC 300
Db 638 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGCGGTGGCTGTATGAGGGCCTGAGC 697
QY 301 AGGAGAAAGCAGAGGAAGTGTGTTTAACTGGGAACCCCTGAGGGGCCCTTCTCATC 360
Db 698 AGGAGAAAGCAGAGGAAGTGTGTTTAACTGGGAACCCCTGAGGGGCCCTTCTCATC 757
QY 361 CGGAGAGCCAGACCAGAGAGGCTCTTACTCTCTGTGAGTCCGCTCAGCCGCCCTGCA 420

Db 758 CGGAGAGCCAGACCAGAGAGGCTTACTCTCTGTGAGTCCGCCCTCAGCCGCCCTGCA 817
QY 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACAATGGCTGTGATCATC 480
Db 818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACAATGGCTGTGATCATC 877
QY 481 TCACCGCGCCTCACCCTTCCCTCACTCCAGGCCCCGTGGACCATTACTCTGAGTGGCG 540
Db 878 TCACCGCGCCTCACCCTTCCCTCACTCCAGGCCCCGTGGACCATTACTCTGAGTGGCG 937
QY 541 GATGACATCTGCTGCTTCAAGAGAGCCCTGTGTCTCTGAGAGGGCTGGCCGCTCCCT 600
Db 938 GATGACATCTGCTGCTTCAAGAGAGCCCTGTGTCTCTGAGAGGGCTGGCCGCTCCCT 997
QY 601 GGCAGGATATACCCCTACCTGTGACTGTGACAGAGACCACTCAACTGAAAGAGCTG 660
Db 998 GGCAGGATATACCCCTACCTGTGACTGTGACAGAGACCACTCAACTGAAAGAGCTG 1057
QY 661 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTCTCAGTGAGGT 720
Db 1058 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTCTCAGTGAGGT 1117
QY 721 CTCGGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGCTGTCTCTTGATGAT 780
Db 1118 CTCGGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGCTGTCTCTTGATGAT 1177
QY 781 GCCTAG 786
Db 1178 GCCTAG 1183

RESULT 7
AAC77202
ID AAC77202 standard; cDNA; 837 BP.

XX
AC AAC77202;

XX
DT 08-FEB-2001 (first entry)

XX
DE Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.

XX
KW Human; open reading frame; ORFX; detection; cytosstatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW thrombosis; contraceptive; ss.

XX
OS Homo sapiens.

XX
PN WO200058473-A2.

XX
PD 05-OCT-2000.

XX
PF 31-MAR-2000; 2000WO-US008621.

XX
PR 31-MAR-1999; 99US-0127607P.

XX
PR 02-APR-1999; 99US-0127636P.

XX
PR 05-APR-1999; 99US-0127728P.

XX
PR 30-MAR-2000; 2000US-00540763.

XX
PA (CURA-) CURAGEN CORP.

XX
PI Shimkets RA, Leach M;

DR WPI; 2000-602362/57.
DR P-PSDB; AAB42993.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 5; Page 4692-4693; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antineumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;

Query Match 94.4%; Score 741.8; DB 3; Length 837;
Best Local Similarity 99.7%; Pred. No. 5.2e-186;
Matches 743; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 42 AAGCTTGAGTTCCTCTGTCACAAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAGCAA 101
DB 3 AAGCTTGAGTTCCTCTGTCACAAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAGCAA 62

QY 102 GGGCCACAGCCGTGGCCCTGGGAGATTTCGCCGGCAGGTGGCCCGGAGCTTCGCTGAG 161
DB 63 GGGCCACAGCCGTGGCCCTGGGAGATTTCGCCGGCAGGTGGCCCGGAGCTTCGCTGAG 122

QY 162 ACTCGGGAGCCATGACCATCGTCTGTGAGATGAGACTGTGTGACCGGTGCTGTGA 221
DB 123 ACTCGGGAGCCATGACCATCGTCTGTGAGATGAGACTGTGTGACCGGTGCTGTGA 182

QY 222 AGTCTCAGGCAGAGATATAACATCCCCAGCCGTCACGTCGCCAAAGTCTTCCATGGGTG 281
DB 183 AGTCTCAGGCAGAGATATAACATCCCCAGCCGTCACGTCGCCAAAGTCTTCCATGGGTG 242

QY 282 GCTGTATGAGGGCCTGAGCAGGGAGAAAGCAGAGAACTGCTGTGTTACTTGGGAACCC 341
DB 243 GCTGTATGAGGGCCTGAGCAGGGAGAAAGCAGAGAACTGCTGTGTTACTTGGGAACCC 302

QY 342 TGGAGGGGCTTCTCTCATCCGGGAGAGCCAGACGAGAGGCTCTTACTCTCTGTCACT 401
DB 303 TGGAGGGGCTTCTCTCATCCGGGAGAGCCAGACGAGAGGCTCTTACTCTCTGTCACT 362

QY 402 CCGCCTCAGCCGCCCTGCATCCTGGAGCCGATCAGACACTACAGGATCCACTGCTTGA 461
DB 363 CCGCCTCAGCCGCCCTGCATCCTGGAGCCGATCAGACACTACAGGATCCACTGCTTGA 422

QY 462 CAATGGCTGGCTGTACATCTACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGTGA 521
DB 423 CAATGGCTGGCTGTACATCTACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGTGA 482

QY 522 CCATTACTCTGAGCTGGCGGATGACATCTGCTGCTTCACTCAAGGAGCCCTGTGCTGCA 581
DB 483 CCATTACTCTGAGCTGGCGGATGACATCTGCTGCTTCACTCAAGGAGCCCTGTGCTGCA 542

QY 582 GAGGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACTGTGACTGTGCAAGAGACACC 641

DB 543 GAGGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACTGCTGACTGTGACAGAGACACC 602

QY 642 ACTCAACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGA 701
DB 603 ACTCAACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGA 662

QY 702 GTCTCTTCTCAGTGAGGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGA 761
DB 663 GTCTCTTCTCAGTGAGGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGA 722

QY 762 GGCTGTCTCTTTGGATGATGCTAG 786
DB 723 GGCTGTCTCTTTGGATGATGCTAG 747

RESULT 8
ABQ99374
ID ABQ99374 standard; cDNA; 1413 BP.
XX
AC ABQ99374;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human coding sequence SEQ ID 107.
XX
KW Human; expressed sequence tag; EST; chromosome 20;
KW hematopoietic disorder; central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; nootropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytosstatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US042950.
XX
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSEQ INC.
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-590824/63.
DR N-PSDB; ABP64788.
XX
PT New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.
XX
PS Claim 1; SEQ ID NO 107; 394pp; English.
XX
CC The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotentia or pluripotentia state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or

CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. haematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

Query Match 85.6%; Score 672.8; DB 6; Length 1413;
Best Local Similarity 93.4%; Pred. No. 1.1e-167;
Matches 734; Conservative 0; Mismatches 2; Indels 50; Gaps 1;

QY 1 ATGGGAAGTCTGCCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
DB 54 ATGGGAGTCTGCCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 113
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 120
DB 114 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 173
QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
DB 174 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 233
QY 181 ATGCTCTCTGAGATGAGACTGTGACCGTGTCTGTAAGTCTCAGGCAGAGAGTAT 240
DB 234 ATGCTCTCTGAGATGAGACTGTGACCGTGTCTGTAAGTCTCAGGCAGAGAGTAT 293
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGACC 300
DB 294 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGACC 353
QY 301 AGGAGAAACGAGAGAACTGTGTTTACCTGGGAACCTTGAGGGGCCCTTCTCATC 360
DB 354 AGGAGAAACGAGAGAACTGTGTTTACCTGGGAACCTTGAGGGGCCCTTCTCATC 413
QY 361 CGGAGAGCCAGACGAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCCCTGCA 420
DB 414 CGGAGAGCCAGACGAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCCCTGCA 473
QY 421 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACATAGGCTGGCTGTACATC 480
DB 474 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACATAGGCTGGCTGTACATC 533
QY 481 TCACCGGCTCACCCTTCCCTCACTCCAGGCCCTGTGAGCACTTACTCTGAGCTGGCG 540
DB 534 TCACCGGCTCACCCTTCCCTCACTCCAGGCCCTGTGAGCACTTACTCTGAGCTGGCG 584
QY 541 GATGACATCTGTGCTGCTACTCAAGAGCCCTGTCTCTGACAGAGGCTGGCCGCTCCCT 600
DB 585 -----GAGGCTGGCCGCTCCCT 603
QY 601 GGCAGGATATACCCCTACTCTGTACTGTGACAGAGGACACCACTCACTGGAAGAGCTG 660
DB 604 GGCAGGATATACCCCTACTCTGTACTGTGACAGAGGACACCACTCACTGGAAGAGCTG 663
QY 661 GACAGCTCCCTCTGTTTCTGAAAGTGGCAGAGGAGAGTCTTCTCTCAGTGAAGGT 720
DB 664 GACAGCTCCCTCTGTTTCTGAAAGTGGCAGAGGAGAGTCTTCTCTCAGTGAAGGT 723
QY 721 CTCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTGGATGAT 780
DB 724 CTCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTGGATGAT 783
QY 781 GCCTAG 786
DB 781 GCCTAG 786

Db 784 GCCTAG 789

RESULT 9
AAL44090
ID AAL44090 standard; cDNA; 737 BP.

XX AAL44090;

XX 03-OCT-2002 (first entry)

DE Mouse MARS short isoform protein coding sequence.

XX Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.

XX Mus sp.

XX Key Location/Qualifiers
FH 1..633
FT CDS /*tag= a
FT /product= "Mouse MARS short isoform protein"

PN WO200242452-A2.

PD 30-MAY-2002.

PF 26-NOV-2001; 2001WO-CA001662.

PR 27-NOV-2000; 2000CA-02324663.

PA (HOSP-) HOSPITAL FOR SICK CHILDREN.

PI Mcglade JC, Loreto MP;

XX MPI; 2002-566564/60.

DR P-PSDB; AAO15458.

XX New isolated modulator of antigen receptor signaling protein or its
PT fragment, useful for treating malignant disorders such as myeloid
PT malignancies, autoimmune disorders and myeloproliferative disorders.

XX Claim 9; Page 77; 110p; English.

XX The invention comprises the amino acid and coding sequences of modulator
CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present cDNA sequence encodes a mouse MARS protein

XX Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;

Query Match 84.4%; Score 663.4; DB 6; Length 737;
Best Local Similarity 93.4%; Pred. No. 2.7e-165;
Matches 735; Conservative 0; Mismatches 1; Indels 51; Gaps 2;

QY 1 ATGGGAGTCTGCCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60

DB 1 ATGGGAGTCTGCCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60

QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG 120

DB 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG 120

QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180

DB 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180

QY 181 ATGCTCTGAGGATGAGACTGGTGAACGGTGTCTGTGAAGTCTCAGGACAGAGATAT 240
Db 181 ATGCTCTGAGGATGAGACTGGTGAACGGTGTCTGTGAAGTCTCAGGACAGAGATAT 240
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGTGTATGAGGGCCTGAGC 300
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGTGTATGAGGGCCTGAGC 300
QY 301 AGGAGAAAGCAGAGAACTGTGTGTTAAGTCTGGAAACCTGGAGGGCCTTCCTCATC 360
Db 301 AGGAGAAAGCAGAGAACTGTGTGTTAAGTCTGGAAACCTGGAGGGCCTTCCTCATC 360
QY 361 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTGTCAATCCGCTCAGCCGCTGCA 420
Db 361 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTGTCAATCCGCTCAGCCGCTGCA 420
QY 421 TCCTGGGACCGGATCAGACACTAAGAGATCCACTGCTTGAATGAGTGGCTGTATCATC 480
Db 421 TCCTGGGACCGGATCAGACACTAAGAGATCCACTGCTTGAATGAGTGGCTGTATCATC 480
QY 481 TCACCGCGCTCACCCTCCCTCACTCCAGGCGCTGTGAGCCATTACTCTGAGTGGCG 540
Db 481 TCACCGCGCTCACCCTCCCTCACTCCAGGCGCTGTGAGCCATTACTCTGAGTGGCG 540
QY 541 GATGACATCTGCTGCTACTCAAGGAGCCCTGTGTCTGAGAGGCGCTGCGCTCCCT 600
Db 541 GATGACATCTGCTGCTACTCAAGGAGCCCTGTGTCTGAGAGGCGCTGCGCTCCCT 600
QY 601 GGCAAGATATATCCCTCACTGTGAGTGTGAGAGACCACTCACTGAAAGAGCTG 660
Db 601 GGCAAGATATATCCCTCACTGTGAGTGTGAGAGACCACTCACTGAAAGAGCTG 660
QY 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTTCTCAAGTGGGT 720
Db 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTTCTCAAGTGGGT 720
QY 721 CTCCGGAGTCCCTGAGCTTCTACATCAGCCTGAATGA-CGAGGCTGTCTTTGATGA 779
Db 721 CTCCGGAGTCCCTGAGCTTCTACATCAGCCTGAATGA-CGAGGCTGTCTTTGATGA 779
QY 780 TGCCTAG 786
Db 780 TGCCTAG 786
QY 731 TGCCTAG 737
Db 731 TGCCTAG 737
RESULT 10
AAS74750
ID AAS74750 standard; cDNA; 2049 BP.
AC AAS74750;
XX 13-FEB-2002 (first entry)
DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #10554.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS Homo sapiens.
PN WO200175067-A2.
XX 11-OCT-2001.
PD 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.
DR P-PSDB; ABG10563.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 10554; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;
Query Match 77.6%; Score 609.8; DB 5; Length 2049;
Best Local Similarity 92.2%; Pred. No. 5.6e-151;
Matches 718; Conservative 0; Mismatches 7; Indels 54; Gaps 5;
QY 1 ATGGAAGTCTGCCAGCAGAGAAATCTTCCCAAGCCCAAGCTTGAGTCTCTGTC 60
Db 965 ATGGAAGTCTGCCAGCAGAGAAATCTTCCCAAGCCCAAGCTTGAGTCTCTGTC 1024
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCCACAGCCCTGCGCTG 120
Db 1025 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCCACAGCCCTGCGCTG 1084
QY 121 GGCAGTTTCCGCGCAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 1085 GGCAGTTTCCGCGCAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 1144
QY 181 ATGCTCTGAGGATGAGACTGTGAGCGGTGTCTGAAGTCTCAGGACAGAGATAT 240
Db 1145 ATGCTCTGAGGATGAGACTGTGAGCGGTGTCTGAAGTCTCAGGACAGAGATAT 1204
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGTGTATGAGGGCCTGAGC 300
Db 1205 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGTGTATGAGGGCCTGAGC 1264
QY 301 AGGAGAAAGCAGAGAACTGTGTGTTAAGTCTGGAAACCTGGAGGGCCTTCCTCATC 360
Db 1265 AGGAGAAAGCAGAGAACTGTGTGTTAAGTCTGGAAACCTGGAGGGCCTTCCTCATC 1324
QY 361 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCAATCCGCTCAGCCGCTGCA 420
Db 1325 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCAATCCGCTCAGCCGCTGCA 1384
QY 421 TCCTGGGACCGGATCAGACACTAAGAGATCCACTGCTTGAATGAGTGGCTGTATCATC 480
Db 1385 TCCTGGGACCGGATCAGACACTAAGAGATCCACTGCTTGAATGAGTGGCTGTATCATC 1444
QY 481 TCACCGCGCTCACCCTCCCTCACTCCAGGCGCTGTGAGCCATTACTCTGAGCTGGCG 540

Db 1445 TCACCGCGCTCACTTCCCTCACTCCAGGCCCTGTGGACCATTAATCTCT----- 1495
QY 541 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGTCTTGACAGAGGCTGCCCTCCCT 600
Db 1496 -----GAGGCTGGCCCTCCCT 1514
QY 601 GGCAAGATATACCCCTAAGCTGTGAGTGTGAGAGACACCACTCAACTGGAAGAAGCTG 660
Db 1515 GGCAAGATATACCCCTAAGCTGTGAGTGTGAGAGACACCACTCAACTGGAAGAAGCTG 1574
QY 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACA-GGGAGAGAGTCTCTTCTAG-TGAGG 718
Db 1575 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAAGGGAGAGTCTCTTCTAGAGAGG 1634
QY 719 GTCTCCGGAGT-CCCTCAGCTTCTACATCAG-CTGAATGACGAGGCTGTCTTTGG 775
Db 1635 GGCTCCGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTTG 1693

RESULT 11

AAL44087
ID AAL44087 standard; cDNA; 1348 BP.
XX AC AAL44087;
XX 03-OCT-2002 (first entry)

Mouse modulator of antigen receptor signalling protein coding sequence.

Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
immunosuppression; myeloproliferative disorder; breast cancer.

Mus sp.

Key Location/Qualifiers
CDS 282..1061
FT /*tag= a
FT /product= "Mouse MARS protein"

WO200242452-A2.

30-MAY-2002.

26-NOV-2001; 2001WO-CA001662.

27-NOV-2000; 2000CA-02324663.

(HOSP-) HOSPITAL FOR SICK CHILDREN.

Mcglade JC, Loreto MP;

WPI; 2002-566564/60.
P-PSDB; AA015456.

New isolated modulator of antigen receptor signaling protein or its
fragment, useful for treating malignant disorders such as myeloid
malignancies, autoimmune disorders and myeloproliferative disorders.

Claim 10; Fig 1A; 110pp; English.

The invention comprises the amino acid and coding sequences of modulator
of antigen receptor signalling (MARS) proteins. The MARS protein is a
putative tumour suppressor gene and exhibits structural and sequence
similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
protein sequences of the invention are useful for the treatment of
myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
disorders, immunosuppression, myeloproliferative disorders and
malignancies related to the de-regulation of tyrosine kinases (e.g.
breast cancer). The present cDNA sequence encodes a mouse MARS protein
Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;

Query Match 68.4%; Score 537.8; DB 6; Length 1348;
Best Local Similarity 81.9%; Pred. No. 5.4e-132;
Matches 646; Conservative 0; Mismatches 137; Indels 6; Gaps 2;

QY 1 ATGGAAGTCTGCCAGCAGAGAAATCTGCAAGCCCAAGCTTGAGTCTCTGTC 60
Db 282 ATGGAAGTCTGCCAGCAGAGAAATCTGCAAGCCCAAGCTTGAGTCTCTGTC 338
QY 61 CAAGCCAGGAGCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 120
Db 339 CCAGACCAAGAACCCGTGTTCATGCAACCAAGAAAGACACAAGGTACAGCTGTGGCCCTG 398
QY 121 GGCAGTTTCCCGCAGAGTGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 399 GGCAGTTTCCCGCAGAGTGGAACAGGCCAGACTATCTCTGAGACTCGGGAGCCGTGACC 458
QY 181 ATGCTCTGAGGATGAGACTGTGTGACCGGTGTCTGAACTCTCAGGCAAGAGTAT 240
Db 459 ATCATCTCTGAGGATGAGATTGTGTGACAGTCCAGTGGAAAGTCTCAGGCAAGAGTAC 518
QY 241 AACATCCCGCAGCGTCCAGCTGAGCCAAAGTCTCCATGGGTGCTGTATGAGGCGCTGAC 300
Db 519 CACATGCCAGTGTGTATGTGCTAAAGTCGCCACGGGTGGCTGTACGAGGGCGCTGACC 578
QY 301 AGGAGAAAGCAGAGAACTGCTGTTGTTACCTGGAACCTTGAGGGGCGCTCTCATC 360
Db 579 CGGAGAAAGCCGAGAACTACTCTGTTACCTGGAACCCCGAGGGCGCTCTCATC 638
QY 361 CGGAGAGCCAGACCAAGAGGCTTACTCTGTGAGTCCGCTCAGCCGCTGCA 420
Db 639 CGGAGAGCCAGACCAAGAGGCTGCTATTCCTGTCGCTCCGATCCGACTCAGCCGCTGCA 698
QY 421 TCTTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACATC 480
Db 699 TCTTGGGACCGGATCAGACACTACAGGATCCAGGATCCAGGCTGTGCAATGGCTGTACATC 758
QY 481 TCACCGCGCTCACTTCCCTCACTCCAGGCCCTGTGTGACCAATTACTGTGAGCTGGCG 540
Db 759 TCACCTCGCTCACTTCCCTCACTCCAGGCCCTGTGTGACCAATTACTGTGAGCTGACA 818
QY 541 GATGACATCTGCTGCTTCAAGAGGCCCTGTGTCTGCAAGGGGCTGGCCGCTCCCT 600
Db 819 GATGACATCTGCTGCTTCAAGAGGCCCTGTGTCTGCAAGGGGCTGGCCGCTCCCT 878
QY 601 GGCAAGATATACCCCTAAGCTGTGAGTGTGAGAGACACCACTCAACTGGAAGAAGCTG 660
Db 879 GGCAAGATATACCCCTAAGCTGTGAGTGTGAGAGACACCACTCAACTGGAAGAAGCTG 938
QY 661 GACAGCTCCCTCTGTTTCTGAG---CTGCCACAGGGAGAGTCTTCTCTCAGTGAG 717
Db 939 GACCGCAGCTCTGTTTCTGAGACACCTGCGAGTGGGAGGATCTCTGCTCAGTGAG 998
QY 718 GGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGAT 777
Db 999 GGGCTCCGAGAGTCCCTCAGTCTCTACATCAGCCCTGGCTGAGGACCCCTTGATGATGCT 1058
QY 778 GATGCTAG 786
Db 1059 TAGCCCTGG 1067

RESULT 12

ABQ98670
ID ABQ98670 standard; DNA; 763 BP.

XX AC ABQ98670;

XX 04-NOV-2002 (first entry)

Human ORF477 coding sequence.

Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulneryary;

KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

OS Homo sapiens.

XX US2002082206-A1.

XX 27-JUN-2002.

XX 30-MAY-2001; 2001US-00867550.

XX 30-MAY-2000; 2000US-0208427P.

XX (LEAC/) LEACH M D.

XX (MEHR/) MEHRABAN F.

XX (CONL/) CONLEY P B.

XX (TOPP/) TOPPER J N.

XX (LAWD/) LAW D.

XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX MPI; 2002-626554/67.

XX P-PSDB; ABP64107.

XX New polypeptide designated ORFX are present in human atherogenic cells

XX and are useful to prevent and treat ORFX-associated disorders including

XX cancer, allergy, wound healing or autoimmune, cardiovascular or

XX inflammatory disease.

XX Claim 2; SEQ ID NO 953; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their

XX coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences

XX were discovered in human atherogenic cells, in particular in platelets

XX and human umbilical vein endothelial cells (HUVEC) and are expressed in

XX many other tissues as well. Atherogenic cells are cells which have the

XX potential to develop atherosclerotic plaques. The ORFX polypeptides and

XX nucleic acids are useful for treating or preventing a pathological

XX condition associated with an ORFX-associated disorder, e.g. cancer,

XX cardiovascular disease, allergy, autoimmune disease, wound healing, blood

XX coagulation disorders or inflammatory disorders. Note: The sequence data

XX for this patent did not form part of the printed specification, but was

XX obtained in electronic format directly from the USPTO web site at

XX seqdata.uspto.gov/sequence.html?DocID=20020082206

XX Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

XX Query Match 60.8%; Score 478; DB 6; Length 763;

XX Best Local Similarity 100.0%; Pred. No. 3e-116;

XX Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAGTGTGCCCCAGACAGAGAAATCTCTGCCAAGCCCCAAGCTTGAGTTCCTCTGTC 60

Db 286 ATGGGAAGTGTGCCCCAGACAGAGAAATCTCTGCCAAGCCCCAAGCTTGAGTTCCTCTGTC 345

QY 61 CAAGGCCAGGACCTGTGACCATGAGACAGAGAGAGCAAGGAGCCACAGCCCTG 120

Db 346 CAAGGCCAGGACCTGTGACCATGAGACAGAGAGAGCAAGGAGCCACAGCCCTG 405

QY 121 GGCAGTTTCCCGGCAAGTGGCCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180

Db 406 GGCAGTTTCCCGGCAAGTGGCCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 465

QY 181 ATGCTCTTGAGGATGAGACTGTGACGCTGTCTGTAAGTCTCAGGACAGAGATAT 240

Db 466 ATGCTCTTGAGGATGAGACTGTGACGCTGTCTGTAAGTCTCAGGACAGAGATAT 525

QY 241 AACATCCCAGCGTCCACAGTGGCCAAAGTCTCCATGGGTGCTGTATGAGGCGCTGACC 300

Db 526 AACATCCCAGCGTCCACAGTGGCCAAAGTCTCCATGGGTGCTGTATGAGGCGCTGACC 585

QY 301 AGGAGAAAGCAGAGGAAGTCTGTTGTTACCTGGGAACCTTGAGGGGCTTCTCATC 360

Db 586 AGGAGAAAGCAGAGGAAGTCTGTTGTTACCTGGGAACCTTGAGGGGCTTCTCATC 645

QY 361 CCGGAGAGCCAGACAGAGAGGCTTACTCTGTGAGTCCGCTCAGCCGCCCTGCA 420

Db 646 CCGGAGAGCCAGACAGAGAGGCTTACTCTGTGAGTCCGCTCAGCCGCCCTGCA 705

QY 421 TCCTGGACCGGATCAGACACTACAGATCCACTGCTTGAACAATGGCTGTACA 478

Db 706 TCCTGGACCGGATCAGACACTACAGATCCACTGCTTGAACAATGGCTGTACA 763

RESULT 13

AA574748

ID AA574748 standard; cDNA; 603 BP.

AA574748;

13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #10552.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB; ABG10561.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity.

XX Claim 1; SEQ ID NO 10552; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

XX and in recombinant production of (II). The polynucleotides are also used

XX in diagnostics as expressed sequence tags for identifying expressed

XX genes. (I) is useful in gene therapy techniques to restore normal

XX activity of (II) or to treat disease states involving (II). (II) is

XX useful for generating antibodies against it, detecting or quantitating a

SQ Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;

Query Match 51.5%; Score 405; DB 5; Length 603;
Best Local Similarity 100.0%; Pred. No. 5.6e-97;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 GGCTTACTCTCTGTGACAGTCCGCTCAGCCGCTGCATCCTGGACCGGATCAGACAC 441
199 GGCTTACTCTCTGTGACAGTCCGCTCAGCCGCTGCATCCTGGACCGGATCAGACAC 258

QY 442 TACAGATCCACTGCTTGACATGCTGGCTGTACATCTCACCGGCTTACCTTCCC 501
259 TACAGATCCACTGCTTGACATGCTGGCTGTACATCTCACCGGCTTACCTTCCC 318

QY 502 TCACTCCAGGCGCTGTGGACCATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTC 561
319 TCACTCCAGGCGCTGTGGACCATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTC 378

QY 562 AAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCTGGCAAGATATACCCCTACCT 621
379 AAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCTGGCAAGATATACCCCTACCT 438

QY 622 GTGACTGTGACAGAGACCACTCACTGAAAGAGCTGACAGCTCCCTCTCTTCTTCT 681
439 GTGACTGTGACAGAGACCACTCACTGAAAGAGCTGACAGCTCCCTCTCTTCTTCT 498

QY 682 GAAGCTGCCACAGGGGAGAGGTCTTCTTCTGAGTGGGTCTCCGGAGTCCCTCAGCTTC 741
499 GAAGCTGCCACAGGGGAGAGGTCTTCTTCTGAGTGGGTCTCCGGAGTCCCTCAGCTTC 558

QY 742 TACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTTAG 786
559 TACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTTAG 603

Db

RESULT 14
ADL63090
ID ADL63090 standard; DNA; 864 BP.

XX ADL63090;
XX 20-MAY-2004 (first entry)

XX Human ovarian cancer DNA marker #21302.
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX Homo sapiens.
XX WO200170979-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009126.
XX 21-MAR-2000; 2000US-0191031P.
XX 25-MAY-2000; 2000US-0207124P.
XX 15-JUN-2000; 2000US-0211940P.
XX 07-JUL-2000; 2000US-0216820P.
XX 25-JUL-2000; 2000US-0220661P.
XX 21-DEC-2000; 2000US-0257672P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Lee J, Lillie J;
XX WPI; 2001-611502/70.
XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.
XX Disclosure; SEQ ID NO 21302; 106bp; English.

XX The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of the patient an antisense oligonucleotide complementary to a marker of the invention. The markers are useful for assessing if a patient is afflicted with ovarian cancer, which involves comparing the level of expression of a marker in a patient sample and a normal level of expression of the expression levels indicates ovarian cancer. The level of expression of a marker corresponds to a secreted protein or to a transcribed polynucleotide or its portion. The level of expression of the marker is assessed by detecting the presence in the sample, a protein or protein fragment corresponding to the marker. The presence of protein or protein fragment is detected using an antibody that specifically binds with the protein or protein fragment. Alternatively, the level of expression of the marker is assessed by detecting the presence of a transcribed polynucleotide which anneals with the marker or anneals with a portion of the polynucleotide comprising the marker, under stringent conditions. The marker is also used for monitoring the progression of ovarian cancer in a patient which involves detecting expression of the marker in a patient sample at a first point in time, repeating the method at a subsequent time and comparing the level of expression. The method is carried out using an ovarian tissue sample. A composition comprising a marker, polypeptide or antibody of the invention is used to treat ovarian cancer. This sequence represents a human ovarian cancer DNA marker of the invention.

SQ Sequence 864 BP; 191 A; 259 C; 242 G; 165 T; 0 U; 7 Other;

Query Match 48.5%; Score 381; DB 5; Length 864;
Best Local Similarity 98.7%; Pred. No. 1.4e-90;
Matches 384; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGACAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTCTGTC 60
450 ATGGGAAGTCTGCCAGACAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTCTGTC 509

QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTG 120
510 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTG 569

QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
570 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 629

QY 181 ATCGTCTTGAGGATGAGACTGTGTGACGGTGTCTGTAAGTCTCAGCAGAGAGTAT 240
630 ATCGTCTTGAGGATGAGACTGTGTGACGGTGTCTGTAAGTCTCAGCAGAGAGTAT 689

QY 241 AACATCCCCAGCGTCCACGTGGCCCAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300
690 AACATCCCCAGCGTCCACGTGGCCCAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 749

QY 301 AGGAGAAAGCAGAGGAAGTGTGTTTACCTGGGAACCTTGAGGGGCTTCTCTATC 360
750 AGGAGAAAGCAGAGGAAGTGTGTTTACCTGGGAACCTTGAGGGGCTTCTCTATC 809

QY 361 CGGAGAGCCAGACCAAGAGAGGCTCTTA 389
810 CGGAGAGCCAGACCAAGAGAGGCTCTTA 838

Db

RESULT 15
ABQ99151
ID ABQ99151 standard; DNA; 875 BP.
XX
AC ABQ99151;

XX	04-NOV-2002	(first entry)
DT		
XX		
DE	Human ORF958	coding sequence.
XX		
KW	Cytostatic;	Cardiant; Anti-allergic; Immunosuppressive; Vulnery;
KW	Antiinflammatory;	gene therapy; human; ORFX; atherogenic; platelet;
KW	human umbilical vein endothelial cell;	HUVEC; atherosclerotic plaque;
KW	cancer; cardiovascular disease;	allergy; autoimmune disease;
KW	wound healing; blood coagulation disorder;	inflammatory disorder; ds.
XX		
OS	Homo sapiens.	
XX		
PN	US2002082206-A1.	
XX		
PD	27-JUN-2002.	
XX		
PF	30-MAY-2001;	2001US-00867550.
XX		
PR	30-MAY-2000;	2000US-0208427P.
XX		
PA	(LEAC/) LEACH M D.	
PA	(MEHR/) MEHRABAN F.	
PA	(CONL/) CONLEY P B.	
PA	(TOPE/) TOPPER J N.	
PA	(LAWD/) LAW D.	
XX		
PI	Leach MD, Mehraban F, Conley PB, Topper JN, Law D;	
XX		
DR	WPI; 2002-626554/67.	
DR	P-PSDB; ABP64588.	
XX		
PT	New polypeptide designated ORFX are present in human atherogenic cells	
PT	and are useful to prevent and treat ORFX-associated disorders including	
PT	cancer, allergy, wound healing or autoimmune, cardiovascular or	
PT	inflammatory disease.	
XX		
PS	Claim 2; SEQ ID NO 1915; 78pp; English.	
XX		
CC	The present invention relates to novel human ORFX polypeptides and their	
CC	coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences	
CC	were discovered in human atherogenic cells, in particular in platelets	
CC	and human umbilical vein endothelial cells (HUVEC) and are expressed in	
CC	many other tissues as well. Atherogenic cells are cells which have the	
CC	potential to develop atherosclerotic plaques. The ORFX polypeptides and	
CC	nucleic acids are useful for treating or preventing a pathological	
CC	condition associated with an ORFX-associated disorder, e.g. cancer,	
CC	cardiovascular disease, allergy, autoimmune disease, wound healing, blood	
CC	coagulation disorders or inflammatory disorders. Note: The sequence data	
CC	for this patent did not form part of the printed specification, but was	
CC	obtained in electronic format directly from the USPTO web site at	
CC	seqdata.uspto.gov/sequence.html?DocID=20020082206	
XX		
SQ	Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;	
XX		
QY	Query Match	43.4%; Score 341; DB 6; Length 875;
QY	Best Local Similarity	100.0%; Pred. No. 5.3e-80;
QY	Matches 341; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DB		
QY	446 GGATCCACTGACCTTGACAATGGCTGCTGTACATCTCACCGCGCTCAACCTTCCCTCAG	505
DB	2 GGATCCACTGACCTTGACAATGGCTGCTGTACATCTCACCGCGCTCAACCTTCCCTCAG	61
QY	506 TCCAGGCCCCGTGTGGACCATTAATCTGTAGCTGGCGGATGACATCTGCTGCTACTCAAGG	565
DB	62 TCCAGGCCCCGTGTGGACCATTAATCTGTAGCTGGCGGATGACATCTGCTGCTACTCAAGG	121
QY	566 AGCCCTGTGTCTGTGACAGAGGCTGGCCGCTCCCTGGCAAGATATACCCCTACCTGTGA	625
DB	122 AGCCCTGTGTCTGTGACAGAGGCTGGCCGCTCCCTGGCAAGATATACCCCTACCTGTGA	181
QY	626 CTGTGACAGAGACCACTCAATGGAAGAGCTGAGACCTCCCTCTGTTTCTGAAG	685
DB		

Db	182	CTGTGCAGAGGACACCACCTCAACTGGAAGAAGCTGACAGCTCCCTCCTGTTTCTGAAG	241
QY	686	CTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACA	745
Db	242	CTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACA	301
QY	746	TCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTTAG	786
Db	302	TCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTTAG	342

Search completed: November 16, 2004, 21:33:29
Job time : 518 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

November 16, 2004, 17:07:09 ; Search time 3766 Seconds
(without alignments)
9869.810 Million cell updates/sec

Title: US-10-043-649-1

Perfect score: 786

Sequence: 1 atgggaagtcgtgccagcag.....tctcttgatgatgcctag 786

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmb1:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	786	100.0	786	6	AX511153	AX511153 Sequence
2	786	100.0	786	6	AX572845	AX572845 Sequence
3	786	100.0	786	9	AF290985	AF290985 Homo sapi
4	786	100.0	786	9	AF326353	AF326353 Homo sapi
5	786	100.0	2415	9	AK025645	AK025645 Homo sapi
6	786	100.0	2538	9	BC042041	BC042041 Homo sapi
7	786	100.0	2567	6	AX452880	AX452880 Sequence
8	785	99.9	2788	6	AX780857	AX780857 Sequence
9	784.4	99.8	1183	6	AX443133	AX443133 Sequence
10	784.4	99.8	1183	6	AX443135	AX443135 Sequence
11	663.4	84.4	737	6	AX511155	AX511155 Sequence
12	663.4	84.4	737	9	AF290986	AF290986 Homo sapi
13	539.4	68.6	1384	10	BC052655	BC052655 Mus muscu
14	537.8	68.4	1348	6	AX511150	AX511150 Sequence
15	537.8	68.4	1348	10	AF287467	AF287467 Mus muscu
16	536.8	68.3	777	6	AX511151	AX511151 Sequence
17	536.2	68.2	1321	10	AF434990	AF434990 Mus muscu
18	441	56.1	441	6	CQ724864	CQ724864 Sequence
19	381	48.5	864	6	CQ414231	CQ414231 Sequence

20	215	27.4	320	6	CQ410243	CQ410243 Sequence
21	214.2	27.3	1571	5	BC056035	BC056035 Xenopus 1
22	160.6	20.4	831	9	CR536537	CR536537 Homo sapi
23	160.6	20.4	2021	9	HSU44403	U44403 Human Str-1
24	160.6	20.4	2670	6	CQ725810	CQ725810 Sequence
25	160.6	20.4	3090	9	BSM808530	BSM808530 Homo sapi
26	160.2	20.4	2109	6	AX428893	AX428893 Sequence
27	159	20.2	1076	9	HSU30473	U30473 Homo sapien
28	159	20.2	1870	9	BC007042	BC007042 Homo sapi
29	159	20.2	2665	6	AX333017	AX333017 Sequence
30	159	20.2	2665	9	D89077	D89077 Homo sapien
31	156.2	19.9	145068	2	HSDJ977B1	AL050318 Human DNA
32	156.2	19.9	145833	2	AC026539	AC026539 Homo sapi
33	141.8	18.0	432	6	CQ051968	CQ051968 Sequence
34	141.8	18.0	432	6	CQ067012	CQ067012 Sequence
35	141.8	18.0	432	6	CQ094063	CQ094063 Sequence
36	141.8	18.0	432	6	CQ132849	CQ132849 Sequence
37	141.8	18.0	432	6	CQ171433	CQ171433 Sequence
38	141.8	18.0	432	6	CQ200546	CQ200546 Sequence
39	141.8	18.0	432	6	CQ216063	CQ216063 Sequence
40	141.8	18.0	432	6	CQ254624	CQ254624 Sequence
41	141.8	18.0	432	6	CQ291709	CQ291709 Sequence
42	141.8	18.0	432	6	CQ328769	CQ328769 Sequence
43	141.8	18.0	448	6	CQ068653	CQ068653 Sequence
44	141.8	18.0	448	6	CQ293385	CQ293385 Sequence
45	139.2	17.7	849	5	AY278230	AY278230 Gallus ga

ALIGNMENTS

RESULT 1	AX511153	786 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	AX511153	Sequence 4 from Patent WO0242452.			
DEFINITION	AX511153				
ACCESSION	AX511153				
VERSION	AX511153.1	GI:23392046			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	Mcglade, J.C. and Loreto, M.P.			
AUTHORS	Adapler gene				
TITLE	Patent: WO 0242452-A 4 30-MAY-2002;				
JOURNAL	The Hospital for Sick Children (CA)				
FEATURES	Location/Qualifiers				
source	1. .786				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN

Query Match	100.0%;	Score 786;	DB 6;	Length 786;
Best local Similarity	100.0%;	Pred. No. 8.5e-195;		
Matches	786;	Conservative 0;	Mismatches 0;	Indels 0;
				Gaps 0;
QY	1	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC	60	
Db	1	ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC	60	
QY	61	CAAGGCCAGGACCTGTGACCATGGAAGAGAGAGCAAGGCCAGCCGTGCCCTG	120	
Db	61	CAAGGCCAGGACCTGTGACCATGGAAGAGAGAGCAAGGCCAGCCGTGCCCTG	120	
QY	121	GGCAGTTTCCCGCAGGTGGCCCGCAGCTGTGCTGAGACTCGGGAGCCATTGACC	180	
Db	121	GGCAGTTTCCCGCAGGTGGCCCGCAGCTGTGCTGAGACTCGGGAGCCATTGACC	180	
QY	181	ATCGTCTGAGATGAGACTGTGTGACCGTGTCTGAAGTCTCAGGACAGAGATAT	240	
Db	181	ATCGTCTGAGATGAGACTGTGTGACCGTGTCTGAAGTCTCAGGACAGAGATAT	240	

QY 241 AACATCCCCAGCGCTCCACGCTGGCCAAAGTCTCCCATGGGTGGCTGTATAGAGGCGCTGAGC 300
DB 241 AACATCCCCAGCGCTCCACGCTGGCCAAAGTCTCCCATGGGTGGCTGTATAGAGGCGCTGAGC 300
QY 301 AGGAGAAAGCAGAGGAAGTCTGTGTGTACTGGGAACCCCTGAGGGGCGCTTCTCATC 360
DB 301 AGGAGAAAGCAGAGGAAGTCTGTGTGTACTGGGAACCCCTGAGGGGCGCTTCTCATC 360
QY 361 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTGTAGTCCGCTCAGCCGCTGCA 420
DB 361 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTGTAGTCCGCTCAGCCGCTGCA 420
QY 421 TCCTGGAGCCGATCAGACACTACAGAGATCCACTGCTTGAATGGCTGGTGTACATC 480
DB 421 TCCTGGAGCCGATCAGACACTACAGAGATCCACTGCTTGAATGGCTGGTGTACATC 480
QY 481 TCACCGCGCCTCACCCTTCCCTCACTCCAGGCGCTGTGGAACCATTAATCTGAGCTGGCG 540
DB 481 TCACCGCGCCTCACCCTTCCCTCACTCCAGGCGCTGTGGAACCATTAATCTGAGCTGGCG 540
QY 541 GATGACATCTGTGCTTCTTCAAGAGAGCCCTGTGTCTGAGAGGGGCTGGCCGCTCCCT 600
DB 541 GATGACATCTGTGCTTCTTCAAGAGAGCCCTGTGTCTGAGAGGGGCTGGCCGCTCCCT 600
QY 601 GCGAAGATATACCCCTTACTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GCGAAGATATACCCCTTACTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GACAGCTCCCTCCTGTTTCTGAAAGCTGCCACAGAGGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GACAGCTCCCTCCTGTTTCTGAAAGCTGCCACAGAGGAGAGAGAGAGAGAGAGAGAG 720
QY 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGAGGCTGTCTCTTTGATGAT 780
DB 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGAGGCTGTCTCTTTGATGAT 780
QY 781 GCCTAG 786
DB 781 GCCTAG 786

RESULT 2
AX572845 786 bp DNA linear PAT 29-NOV-2002
LOCUS AX572845
DEFINITION Sequence 1 from Patent WO02055707.
ACCESSION AX572845
VERSION AX572845.1 GI:26004935
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Holland,S.J., Mendenhall,M.K., Pardo,J., Spencer,C., Fu,A.C.,
Luo,Y., Payan,D.G., Mancebo,H.S., Wu,J., Zhou,X., Shen,M.,
Liao,X.C. and Sheng,N.
TITLE Cloning of an inhibitor of antigen-receptor signaling by a
retroviral-based functional screen
JOURNAL Patent: WO 02055707-A 1 18-JUL-2002;
Rigel Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..786
/note="unamed protein product"

CDS
/codon_start=1
/protein_id="CAD58542.1"
/db_xref="GI:26004936"
/translation="MGSLPSRSLPSLSSVQGGPVTWEAERSKATAVALGSFP
AGGPAELSLRIGEPLTIVSEBDGWTVLSEVSGREYNIPSVHAKVSHGWLVEGLSR

ORIGIN
Query Match 100.0%; Score 786; DB 6; Length 786;
Best Local Similarity 100.0%; Pred. No. 8.5e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGCAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAATTCTCTGTC 60
DB 1 ATGGGAAGTCTGCCAGCAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAATTCTCTGTC 60
QY 61 CAAGCCAGGAGCTGTGATCCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGCGCTG 120
DB 61 CAAGCCAGGAGCTGTGATCCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGCGCTG 120
QY 121 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTGCTGAGACTGGGGAGCCATTGACC 180
DB 121 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTGCTGAGACTGGGGAGCCATTGACC 180
QY 181 ATCGTCTTGAGATGAGACTGTGTGACGCTGTGTCTGAAGTCTCAGCAGAGAGATAT 240
DB 181 ATCGTCTTGAGATGAGACTGTGTGACGCTGTGTCTGAAGTCTCAGCAGAGAGATAT 240
QY 241 AACATCCCCAGCGCTCCACGCTGGCCAAAGTCTCCCATGGGTGGCTGTATAGAGGCGCTGAGC 300
DB 241 AACATCCCCAGCGCTCCACGCTGGCCAAAGTCTCCCATGGGTGGCTGTATAGAGGCGCTGAGC 300
QY 301 AGGAGAAAGCAGAGGAAGTCTGTGTTACCTGGGAACCTGGAGGGGCGCTTCTCATC 360
DB 301 AGGAGAAAGCAGAGGAAGTCTGTGTTACCTGGGAACCTGGAGGGGCGCTTCTCATC 360
QY 361 CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTCAATCCGCTCAGCCGCTGCA 420
DB 361 CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTCAATCCGCTCAGCCGCTGCA 420
QY 421 TCCTGGAGCCGATCAGACACTACAGAGTCCACTGCTTGAACAATGGCTGGTGTACATC 480
DB 421 TCCTGGAGCCGATCAGACACTACAGAGTCCACTGCTTGAACAATGGCTGGTGTACATC 480
QY 481 TCACCGCGCCTCACCCTTCCCTCACTCCAGGCGCTGTGGAACCATTAATCTGAGCTGGCG 540
DB 481 TCACCGCGCCTCACCCTTCCCTCACTCCAGGCGCTGTGGAACCATTAATCTGAGCTGGCG 540
QY 541 GATGACATCTGTGCTTCTTCAAGAGAGCCCTGTGTCTGCTGAGAGGGCTGGCCGCTCCCT 600
DB 541 GATGACATCTGTGCTTCTTCAAGAGAGCCCTGTGTCTGCTGAGAGGGCTGGCCGCTCCCT 600
QY 601 GCGAAGATATACCCCTTACTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GCGAAGATATACCCCTTACTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GACAGCTCCCTCCTGTTTCTGAAAGCTGCCACAGAGGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GACAGCTCCCTCCTGTTTCTGAAAGCTGCCACAGGAGAGAGAGAGAGAGAGAGAG 720
QY 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGAGGCTGTCTTTGATGAT 780
DB 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGAGGCTGTCTTTGATGAT 780
QY 781 GCCTAG 786
DB 781 GCCTAG 786

RESULT 3
AF290985 786 bp mRNA linear PRI 21-JAN-2003
LOCUS AF290985
DEFINITION Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF290985
VERSION AF290985.1 GI:17351920
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 786)
TITLE Loreto, M.P. and McGlade, C.J.
JOURNAL Cloning and characterization of human Src-1-like adaptor protein 2
MEDLINE Oncogene 22 (2), 266-273 (2003)
PUBMED 22415750
REFERENCE 2 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
source
1. .786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1. .786
/note="SLAP-2"
/codon_start=1
/product="Src-1-like adaptor protein-2"
/protein_id="AAL38197.1"
/db_xref="GI:17351921"
/translation="MSGSLPSRRKSLSPSLSSVQGGPVTMEERSKATAVALGSFP
AGGPAELSLRLGEPLTIVSEDDGDMWTIVSEVSGREYNIPSVHVAKVSHGLYEGLSRE
KAEELLLPGNPGAFILIRISQTRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYI
SPRLTFPSIALVDHYSELADICCLIKPEPCVLQKAGPLPGKDIPLPVTQRTPLNMWK
ELDSLSLFSSEATGEESLSLSEGLRESLSFYISLNDENVSLDDA"
ORIGIN
Query Match 100.0%; Score 786; DB 9; Length 786;
Best Local Similarity 100.0%; Pred. No. 8.5e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
Db 1 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGCAAGCAAGGCCACAGCCGTGGCCCTG 120
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGCAAGCAAGGCCACAGCCGTGGCCCTG 120
QY 121 GGCAGTTTCCCGCAGAGTGGCCCCGCGAGCTGTCCTGAGACTCGGGAGCCATTGACC 180
Db 121 GGCAGTTTCCCGCAGAGTGGCCCCGCGAGCTGTCCTGAGACTCGGGAGCCATTGACC 180
QY 181 ATGCTCTGTGAGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT 240
Db 181 ATGCTCTGTGAGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT 240
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCATGGGTGCTGTATGAGGGCCTGAGC 300
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCATGGGTGCTGTATGAGGGCCTGAGC 300
QY 301 AGGGAAGACAGAGAACTGCTGTGTGTAACCTGGGAACCCCTGAGGGGCTTCTCATC 360
Db 301 AGGGAAGACAGAGAACTGCTGTGTGTAACCTGGGAACCCCTGAGGGGCTTCTCATC 360
QY 361 CGGAGAGCCAGACCCAGAGAGGCTCTACTCTCTGTCACTCCGCTCAAGCCGCTTCA 420
Db 361 CGGAGAGCCAGACCCAGAGAGGCTCTACTCTCTGTCACTCCGCTCAAGCCGCTTCA 420
QY 421 TCCTGGACCGGATCAGACACTACAGATCCACTGCCTTGACCAATGGCTGGCTGTACATC 480
Db 421 TCCTGGACCGGATCAGACACTACAGATCCACTGCCTTGACCAATGGCTGGCTGTACATC 480
QY 481 TCACCGGCTCACTTCCCTCACTCCAGGCCCTGTGTGACCATTAATCTGAGCTGGCG 540

Db 481 TCACCGGCTCACTTCCCTCACTCCAGGCCCTGTGTGACCATTAATCTGAGCTGGCG 540
QY 541 GATGACATCTGCTGCTTACTCAAGAGGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCT 600
Db 541 GATGACATCTGCTGCTTACTCAAGAGGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCT 600
QY 601 GGCAAGATATACCCCTTACTCTGTACTGTGACAGAGACCACTCAACTGGAAGAGCTG 660
Db 601 GGCAAGATATACCCCTTACTCTGTACTGTGACAGAGACCACTCAACTGGAAGAGCTG 660
QY 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT 720
Db 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT 720
QY 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGATGAT 780
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGATGAT 780
QY 781 GCCTAG 786
Db 781 GCCTAG 786
RESULT 4
AF326353 786 bp mRNA linear PRI 08-NOV-2001
LOCUS Homo sapiens Src-1-like adapter protein-2 mRNA, complete cds.
DEFINITION AF326353
ACCESSION AF326353
VERSION AF326353.1 GI:16797891
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 786)
Holland, S.J., Liao, X.C., Mendenhall, M.K., Zhou, X., Pardo, J.,
Chu, P., Spencer, C., Fu, A.C., Sheng, N., Yu, P., Pali, E., Nagin, A.,
Shen, M., Xu, S., Chan, E., Wu, X., Li, C., Woltschlag, M.,
Aversa, G., Kolbinger, F., Bennett, M.K., Molineaux, S., Luo, Y.,
Payan, D.G., Mancebo, H.S.Y. and Wu, J.
TITLE Functional cloning of Src-1-like Adapter Protein-2 (SLAP-2), a Novel
Inhibitor of Antigen Receptor Signaling
JOURNAL J. Exp. Med. 194 (9), 1263-1276 (2001)
MEDLINE 21553259
PUBMED 11696592
REFERENCE 2 (bases 1 to 786)
Holland, S.J., Mendenhall, M.K., Zhou, X., Spencer, C., Pardo, J.,
Fu, A.C., Sheng, N., Shen, M., Liao, C., Luo, Y., Payan, D.G.,
Mancebo, H.S.Y. and Wu, J.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand
Avenue, South San Francisco, CA 94080, USA
FEATURES
source
1. .786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
1. .786
/note="SLAP2; SH2 and SH3 domain-containing adaptor"
/codon_start=1
/product="Src-1-like adaptor protein-2"
/protein_id="AAL29204.1"
/db_xref="GI:16797892"
/translation="MSGSLPSRRKSLSPSLSSVQGGPVTMEERSKATAVALGSFP
AGGPAELSLRLGEPLTIVSEDDGDMWTIVSEVSGREYNIPSVHVAKVSHGLYEGLSRE
KAEELLLPGNPGAFILIRISQTRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYI
SPRLTFPSIALVDHYSELADICCLIKPEPCVLQKAGPLPGKDIPLPVTQRTPLNMWK
ELDSLSLFSSEATGEESLSLSEGLRESLSFYISLNDENVSLDDA"
ORIGIN
Query Match 100.0%; Score 786; DB 9; Length 786;

Best Local Similarity 100.0%; Pred. No. 8.5e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGAAGTCTGCCAGACAGAGAAAATCTTTCGCCAAGCCCAAGCTTGAGTTCTCTGTC	60
Db	1	ATGGGAAGTCTGCCAGACAGAGAAAATCTTTCGCCAAGCCCAAGCTTGAGTTCTCTGTC	60
QY	61	CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG	120
Db	61	CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG	120
QY	121	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC	180
Db	121	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC	180
QY	181	ATCGTCTTGAGGATGAGACTGTGTGACGGTGTCTGAACTCTCAGGACAGAGATAT	240
Db	181	ATCGTCTTGAGGATGAGACTGTGTGACGGTGTCTGAACTCTCAGGACAGAGATAT	240
QY	241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	300
Db	241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	300
QY	301	AGGGAAGAAAGCAGAGAGAACTGTGTGTACCTGGGAACCCCTGAGGGCCCTCTCATC	360
Db	301	AGGGAAGAAAGCAGAGAGAACTGTGTGTACCTGGGAACCCCTGAGGGCCCTCTCATC	360
QY	361	CGGAGAGCCAGACCAAGAGAGGCTCTTACTCTCTGTAGTCCGCTCAGCCGCCCTGCA	420
Db	361	CGGAGAGCCAGACCAAGAGAGGCTCTTACTCTCTGTAGTCCGCTCAGCCGCCCTGCA	420
QY	421	TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGAACAATGGCTGGCTGTACATC	480
Db	421	TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGAACAATGGCTGGCTGTACATC	480
QY	481	TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTGTGACCACTTACTCTGAGCTGGCG	540
Db	481	TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTGTGACCACTTACTCTGAGCTGGCG	540
QY	541	GATGACATCTGTGCTGCTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCT	600
Db	541	GATGACATCTGTGCTGCTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCT	600
QY	601	GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCACTGGAAGAGCTG	660
Db	601	GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCACTGGAAGAGCTG	660
QY	661	GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT	720
Db	661	GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAAGGT	720
QY	721	CTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACAGAGGCTGTCTTTGATGAT	780
Db	721	CTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACAGAGGCTGTCTTTGATGAT	780
QY	781	GCCTAG 786	
Db	781	GCCTAG 786	

RESULT 5

AK025645 2415 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ21992 fis, clone HEP06554.
DEFINITION AK025645
ACCESSION AK025645
VERSION AK025645.1 GI:10438227
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,

TITLE Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
JOURNAL NEDO human cDNA sequencing project
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 2415)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.

TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:ficdha@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES Location/Qualifiers
source 1..2415

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP06554"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_id="HEP"
/note="Cloning vector pME18FL3"
61..846
/note="unnamed protein product"
/protein_id="BAB15201.1"
/codon_start=1
/db_xref="GI:10438228"
/translation="MGSLSRKRSLPSPSLSSVQGGPVTMEARSKATVAALGSP
AGPAELSLRGEPLTIVSEGDWTVLSEVSGREYNIPSHVAKVSHGMYEGLSRE
KAEELLPGNPGAFILRESQTRGYSLSVRLSPASWDRIHRHICLDNGMLYI
SPRLTFPSLQALVDHYSELADDCCLKEPCVLQRAGPLPKDIPLPVTVQRPDLNWK
ELDSLLFSEATGEESLSLSEGLRESLSFYISLNDVAVSLDDA"

ORIGIN

Query Match 100.0%; Score 786; DB 9; Length 2415;
Best Local Similarity 100.0%; Pred. No. 8.2e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGAAGTCTGCCAGACAGAGAAAATCTTTCGCCAAGCCCAAGCTTGAGTTCTCTGTC	60
Db	61	ATGGGAAGTCTGCCAGACAGAGAAAATCTTTCGCCAAGCCCAAGCTTGAGTTCTCTGTC	120
QY	61	CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG	120
Db	121	CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG	180
QY	121	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC	180
Db	181	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC	240
QY	181	ATCGTCTTGAGGATGAGACTGTGTGACGGTGTGTGAACTCTCAGGACAGAGATAT	240
Db	241	ATCGTCTTGAGGATGAGACTGTGTGACGGTGTGTGAACTCTCAGGACAGAGATAT	300
QY	241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	300
Db	301	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	360
QY	301	AGGGAAGAAAGCAGAGAGAACTGTGTGTTAACCCTGGAAACCCCTGAGGGCCCTCTCATC	360
Db	361	AGGGAAGAAAGCAGAGAGAACTGTGTGTTAACCCTGGAAACCCCTGAGGGCCCTCTCATC	420
QY	361	CGGAGAGCCAGACCAAGAGAGGCTCTTACTCTGTGACGCGCTCAGCCGCCCTGCA	420
Db	421	CGGAGAGCCAGACCAAGAGAGGCTCTTACTCTGTGACGCGCTCAGCCGCCCTGCA	480

QY 421 TCCTGGACCGGATCAGACACTACAGGATCCACTGCGCTTGACATGCGCTGATCATC 480
Db 481 TCCTGGACCGGATCAGACACTACAGGATCCACTGCGCTTGACATGCGCTGATCATC 540
QY 481 TCACCGCGGCTCACCCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 540
Db 541 TCACCGCGGCTCACCCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 600
QY 541 GATGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 601 GATGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 601 GGCAAGGATATATACCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 661 GGCAAGGATATATACCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 661 GACAGCTCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 721 GACAGCTCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 721 CTCCGGAGTCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 781 CTCCGGAGTCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 781 GCCTAG 786
Db 841 GCCTAG 846

RESULT 6
LOCUS BC042041 2538 bp mRNA linear PRI 30-JUN-2004
DEFINITION Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (CDNA
clone MGC:49845 IMAGE:4429896), complete cds.
ACCESSION BC042041
VERSION BC042041 GI:27469842
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 2538)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buelow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
JOURNAL
PUBMED 12477932
REFERENCE 2 (bases 1 to 2538)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (23-DEC-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) md@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

FEATURES
source
1. .2538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:49845 IMAGE:4429896"
/tissue_type="Prostate, adenocarcinoma."
/clone_id="NIH MGC_91"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2538
/gene="SLA2"
/note="synonyms: SLAP-2, MGC49845, FLJ21992, SLAP2,
C20orf156"
/db_xref="LocusID:84174"
/db_xref="MIM:606577"
363. .1148
/gene="SLA2"
/codon_start=1
/product="Src-like-adaptor 2, isoform a"
/protein_id="AAH42041.1"
/db_xref="GI:27469843"
/db_xref="MIM:606577"
/db_xref="LocusID:84174"
/translation="MGSLSRSLSPSLSSSVQGGPVMEARSKATAVALGSP
AGGPAELSLRGEPLTIVSEDDWMTVLSEVSGREYNIPSVYAKVSHGWLVEGLSRE
KAEELLLPNNPGAFILRESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYI
SPRLTFPSLOALVDHYSELADICLLKEPCVLQGRAGPLPGKIDPLPVTVQRTPLNWK
ELDSLSLSEATGEESLSLEGRSLSFYISLNDFAVSLDDA"

ORIGIN
Query Match 100.0%; Score 786; DB 9; Length 2538;
Best Local Similarity 100.0%; Pred. No. 8.2e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGAGTCTGCCCCAGCAGAGAAATCTCTGCCAAGCCCCAAGCTTGAAGTCTCTGTC 60
Db 363 ATGGAGAGTCTGCCCCAGCAGAGAAATCTCTGCCAAGCCCCAAGCTTGAAGTCTCTGTC 422
QY 61 CAAGGCCAGGACCTGTGACCATGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 423 CAAGGCCAGGACCTGTGACCATGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
QY 121 GGCAAGTTCCTCCGAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 483 GGCAAGTTCCTCCGAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 542
QY 181 ATGCTCTGAGGATGAGACTGTGTGACCGGTGTGTCTGAAGTCTCAGGCAGAGAGATAT 240
Db 543 ATGCTCTGAGGATGAGACTGTGTGACCGGTGTGTCTGAAGTCTCAGGCAGAGAGATAT 602
QY 241 AACATCCCAAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 300
Db 603 AACATCCCAAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 662
QY 301 AGGAGAAAGCAGAGAGAACTGCTGTTTACTCTGGAAACCTTGAGAGGGCCCTTCTCATC 360

```
|||||
Db 663 AGGAGAAAGCAGAGAACTGCTGTTGTTACTGGAAACCTTGAGGGGCTTCCTCATC 722
QY 361 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTAGTCCGCCCTGAGCCCTGCA 420
Db 723 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTAGTCCGCCCTGAGCCCTGCA 782
QY 421 TCCTGGAGCCGATCAGACACTACAGGATCCACTGCTTGACAATGGCTGGCTGTACATC 480
Db 783 TCCTGGAGCCGATCAGACACTACAGGATCCACTGCTTGACAATGGCTGGCTGTACATC 842
QY 481 TCACCGGCTCCTACCTTCCCTCACTCCAGGCCCTGGTGACCATTAATTAATTAATTAATTA 540
Db 843 TCACCGGCTCCTACCTTCCCTCACTCCAGGCCCTGGTGACCATTAATTAATTAATTAATTA 902
QY 541 GATGACATCTGCTGCTTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCT 600
Db 903 GATGACATCTGCTGCTTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCT 962
QY 601 GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACACCCTCACTCACTGAAAGAGCTG 660
Db 963 GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACACCCTCACTCACTGAAAGAGCTG 1022
QY 661 GACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCTCACTGAGGGT 720
Db 1023 GACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCTCACTGAGGGT 1082
QY 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT 780
Db 1083 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT 1142
QY 781 GCCTAG 786
Db 1143 GCCTAG 1148
```

```
RESULT 7
AX452880 2567 bp DNA linear PAT 06-JUL-2002
LOCUS AX452880
DEFINITION Sequence 1 from Patent WO0242457.
ACCESSION AX452880
VERSION AX452880.1 GI:21712520
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and
Kanner, S.B.
TITLE Cloning and expression of human slap-2: a novel sh2/sh3
domain-containing human slap homologue having immune cell-specific
expression
JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;
Bristol-Myers Squibb Co. (US)
FEATURES
source 1..2567
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
```

```
Query Match 100.0%; Score 786; DB 6; Length 2567;
Best Local Similarity 100.0%; Pred. No. 8.2e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGGAGAGTCTGCCAGCAGAGAAATCTCTCCAAAGCCCAAGCTTGAGTCTCTGTG 60
Db 415 ATGGAGAGTCTGCCAGCAGAGAAATCTCTCCAAAGCCCAAGCTTGAGTCTCTGTG 474
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 120
Db 475 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 534
```

```
QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 535 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 594
QY 181 ATGCTCTGAGAGATGAGAGACTGTGAGCGGTGCTGTCTGAAGTCTCAGGACAGAGATAT 240
Db 595 ATGCTCTGAGAGATGAGAGACTGTGAGCGGTGCTGTCTGAAGTCTCAGGACAGAGATAT 654
QY 241 AACATCCCAAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATAGGGCCCTGAGC 300
Db 655 AACATCCCAAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATAGGGCCCTGAGC 714
QY 301 AGGAGAAAGCAGAGAACTGCTGTTGTTAAGTCCAGGAAACCTTGAGGGGCTTCCTCATC 360
Db 715 AGGAGAAAGCAGAGAACTGCTGTTGTTAAGTCCAGGAAACCTTGAGGGGCTTCCTCATC 774
QY 361 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTCAAGTCCGCCCTCAGCCGCTGCA 420
Db 775 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTCAAGTCCGCCCTCAGCCGCTGCA 834
QY 421 TCCTGGAGCCGATCAGACACTACAGGATCCACTGCTTGACAATGGCTGGCTGTACATC 480
Db 835 TCCTGGAGCCGATCAGACACTACAGGATCCACTGCTTGACAATGGCTGGCTGTACATC 894
QY 481 TCACCGGCTCCTACCTTCCCTCACTCCAGGCCCTGGTGACCATTAATTAATTAATTAATTA 540
Db 895 TCACCGGCTCCTACCTTCCCTCACTCCAGGCCCTGGTGACCATTAATTAATTAATTAATTA 954
QY 541 GATGACATCTGCTGCTTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCT 600
Db 955 GATGACATCTGCTGCTTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCT 1014
QY 601 GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACACCCTCACTGAAAGAGCTG 660
Db 1015 GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACACCCTCACTGAAAGAGCTG 1074
QY 661 GACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAAGGT 720
Db 1075 GACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAAGGT 1134
QY 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT 780
Db 1135 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT 1194
QY 781 GCCTAG 786
Db 1195 GCCTAG 1200
```

```
RESULT 8
AX780857 2788 bp DNA linear PAT 14-JUL-2003
LOCUS AX780857
DEFINITION Sequence 3014 from Patent WO03039443.
ACCESSION AX780857
VERSION AX780857.1 GI:32697851
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.
TITLE Novel genetic markers for leukemias
JOURNAL Patent: WO 03039443-A 3014 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE);
Ludwig-Maximilian-Universitaet Muenchen (DE);
PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
FEATURES
source 1..2788
/organism="Homo sapiens"
/mol_type="unassigned DNA"
```


ORIGIN /db_xref="taxon:9606"

Query Match 99.9%; Score 785; DB 6; Length 2788;
Best Local Similarity 99.9%; Pred. No. 1.5e-194;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGCAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCCTGTG 60
Db 387 ATGGGAAGTCTGCCAGCAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCCTGTG 446
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 120
Db 447 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 506
QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 507 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 566
QY 181 ATGCTCTCTGAGATGAGACTGTGTGACCGGTGTGTGAAAGTCTCAGGCAGAGAGTAT 240
Db 567 ATGCTCTCTGAGATGAGACTGTGTGACCGGTGTGTGAAAGTCTCAGGCAGAGAGTAT 626
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC 300
Db 627 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC 686
QY 301 AGGAGAAAGCAGAGAACTGCTGTGTATCCTGGGAAACCTGGAGGGGCTTCTCATC 360
Db 687 AGGAGAAAGCAGAGAACTGCTGTGTATCCTGGGAAACCTGGAGGGGCTTCTCATC 746
QY 361 CGGAGAGCCAGACAGAGAGAGGCTTCTTACTCTCTGTCAGTCCGCTCAGCCGCCCTGCA 420
Db 747 CGGAGAGCCAGACAGAGAGAGGCTTCTTACTCTCTGTCAGTCCGCTCAGCCGCCCTGCA 806
QY 421 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACAATGGCTGGCTGTATCATC 480
Db 807 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACAATGGCTGGCTGTATCATC 866
QY 481 TCACCGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGTGAGACCATTAATGAGCTGGCG 540
Db 867 TCACCGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGTGAGACCATTAATGAGCTGGCG 926
QY 541 GATGACATCTGCTGCTTACTCAAGAGAGCGCTGTGTCTGACAGAGGGCTGCGCTCCCT 600
Db 927 GATGACATCTGCTGCTTACTCAAGAGAGCGCTGTGTCTGACAGAGGGCTGCGCTCCCT 986
QY 601 GGCAAGATATACCCCTTACTGCTGTGACTGTGAGAGACCACTCAACTGGAAGAAGCTG 660
Db 987 GGCAAGATATACCCCTTACTGCTGTGACTGTGAGAGACCACTCAACTGGAAGAAGCTG 1046
QY 661 GACAGCTCCCTCTCTGTTTCTGAAGCTGACACAGGGAGAGTCTTCTCAGTGAAGGT 720
Db 1047 GACAGCTCCCTCTCTGTTTCTGAAGCTGACACAGGGAGAGTCTTCTCAGTGAAGGT 1106
QY 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTTGGATGAT 780
Db 1107 CTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTTGGATGAT 1166
QY 781 GCCTAG 786
Db 1167 GCCTAG 1172

RESULT 9
LOCUS AX443133 1183 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 74 from Patent W00216599.
ACCESSION AX443133
VERSION AX443133.1 GI:21690555
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
1 Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.
TITLE
JOURNAL
Proteins and nucleic acids encoding same
Patent: WO 0216599-A 74 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
location/Qualifiers
1. .1183
source /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 784.4; DB 6; Length 1183;
Best Local Similarity 99.9%; Pred. No. 2.2e-194;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGCAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCCTGTG 60
Db 398 ATGGGAAGTCTGCCAGCAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCCTGTG 457
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 120
Db 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 517
QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 577
QY 181 ATGCTCTCTGAGATGAGACTGTGTGACCGGTGTGTGAAAGTCTCAGGCAGAGAGTAT 240
Db 578 ATGCTCTCTGAGATGAGACTGTGTGACCGGTGTGTGAAAGTCTCAGGCAGAGAGTAT 637
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC 300
Db 638 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC 697
QY 301 AGGAGAAAGCAGAGAACTGCTGTGTATCCTGGGAAACCTGAGAGGGCTTCTCATC 360
Db 698 AGGAGAAAGCAGAGAACTGCTGTGTATCCTGGGAAACCTGAGAGGGCTTCTCATC 757
QY 361 CGGAGAGCCAGACAGAGAGGCTTCTTACTCTGTGACGTCCGCTCAGCCGCCCTGCA 420
Db 758 CGGAGAGCCAGACAGAGAGGCTTCTTACTCTGTGACGTCCGCTCAGCCGCCCTGCA 817
QY 421 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACAATGGCTGGCTGTATCATC 480
Db 818 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACAATGGCTGGCTGTATCATC 877
QY 481 TCACCGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGTGAGACCATTAATGAGCTGGCG 540
Db 878 TCACCGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGTGAGACCATTAATGAGCTGGCG 937
QY 541 GATGACATCTGCTGCTTACTCAAGAGAGCGCTGTGTCTGACAGAGGGCTGCGCTCCCT 600
Db 938 GATGACATCTGCTGCTTACTCAAGAGAGCGCTGTGTCTGACAGAGGGCTGCGCTCCCT 997
QY 601 GGCAAGATATACCCCTTACTGCTGTGACTGTGAGAGACCACTCAACTGGAAGAAGCTG 660
Db 998 GGCAAGATATACCCCTTACTGCTGTGACTGTGAGAGACCACTCAACTGGAAGAAGCTG 1057
QY 661 GACAGCTCCCTCTCTGTTTCTGAAGCTGACACAGGGAGAGTCTTCTCAGTGAAGGT 720
Db 1058 GACAGCTCCCTCTCTGTTTCTGAAGCTGACACAGGGAGAGTCTTCTCAGTGAAGGT 1117
QY 721 CTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTTGGATGAT 780
Db 1118 CTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTTGGATGAT 1177
QY 781 GCCTAG 786

Db 1178 GCCTAG 1183

RESULT 10
AX443135/c 1183 bp DNA linear PAT 02-JUL-2002
LOCUS AX443135
DEFINITION Sequence 76 from Patent WO0216599.
ACCESSION AX443135
VERSION AX443135.1 GI:21690556
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES
source 1..1183
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 784.4; DB 6; Length 1183;
Best Local Similarity 99.9%; Pred. No. 2.2e-194;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
Db 786 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 727

QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 120
Db 726 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 667

QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
Db 666 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 607

QY 181 ATCGTCTGAGGATGAGACTGTGAGACGGTGTCTGTAAGTCTCAGGCAGAGAGAT 240
Db 606 ATCGTCTGAGGATGAGACTGTGAGACGGTGTCTGTAAGTCTCAGGCAGAGAGAT 547

QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGCCCTGAGC 300
Db 546 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGCCCTGAGC 487

QY 301 AGGAGAAAGCAGAGAACTGCTGTGTACTCTGGAAACCTGAGGGGCTTCTCTCATC 360
Db 486 AGGAGAAAGCAGAGAACTGCTGTGTACTCTGGAAACCTGAGGGGCTTCTCTCATC 427

QY 361 CGGAGAGCCAGACCAAGAGGCTTACTCTCTGTCAAGTCCGCTCAGCCGCCCTGCA 420
Db 426 CGGAGAGCCAGACCAAGAGGCTTACTCTCTGTCAAGTCCGCTCAGCCGCCCTGCA 367

QY 421 TCCTGGAGCCGATCAGACACTACAGATCACTGCTTGAACAATGGCTGGCTGTACATC 480
Db 366 TCCTGGAGCCGATCAGACACTACAGATCACTGCTTGAACAATGGCTGGCTGTACATC 307

QY 481 TCACCCGCGCTCACTTCCCTCACTCCAGGCCCTGGTGAGCCATTACTTGAAGCTGGG 540
Db 306 TCACCCGCGCTCACTTCCCTCACTCCAGGCCCTGGTGAGCCATTACTTGAAGCTGGG 247

QY 541 GATGACATCTGCTGCTACTCAAGGAGCCCTGTCTCTGACAGAGGGCTGGCCCTCCCT 600
Db 246 GATGACATCTGCTGCTACTCAAGGAGCCCTGTCTCTGACAGAGGGCTGGCCCTCCCT 187

QY 601 GGCAGGATATACCCCTACTCTGTGACTGTGACAGAGACACCACCTCACTGGAAGAGCTG 660
Db 186 GGCAGGATATACCCCTACTCTGTGACTGTGACAGAGACACCACCTCACTGGAAGAGCTG 127

QY 661 GACAGCTCCCTCTGTTTCTGAAAGCTGCCACAGGGGAGAGTCTCTCACTGAGGGT 720
Db 126 GACAGCTCCCTCTGTTTCTGAAAGCTGCCACAGGGGAGAGTCTCTCACTGAGGGT 67

QY 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAAATGACGAGGCTGTCTTTGATGAT 780
Db 66 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAAATGACGAGGCTGTCTTTGATGAT 7

QY 781 GCCTAG 786
Db 6 GCCTAG 1

RESULT 11
AX511155 737 bp DNA linear PAT 27-SEP-2002
LOCUS AX511155
DEFINITION Sequence 6 from Patent WO0242452.
ACCESSION AX511155
VERSION AX511155.1 GI:23392047
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Meglade,J.C. and Loreto,M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 6 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source 1..737
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 84.4%; Score 663.4; DB 6; Length 737;
Best Local Similarity 93.4%; Pred. No. 1.1e-162;
Matches 735; Conservative 0; Mismatches 1; Indels 51; Gaps 2;

QY 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60

QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 120
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 120

QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
Db 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180

QY 181 ATCGTCTGAGGATGAGACTGTGAGACGGTGTCTGTAAGTCTCAGGCAGAGAGAT 240
Db 181 ATCGTCTGAGGATGAGACTGTGAGACGGTGTCTGTAAGTCTCAGGCAGAGAGAT 240

QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGCCCTGAGC 300
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGCCCTGAGC 300

QY 301 AGGAGAAAGCAGAGAACTGCTGTGTACTCTGGAAACCTGAGGGGCTTCTCTCATC 360
Db 301 AGGAGAAAGCAGAGAACTGCTGTGTACTCTGGAAACCTGAGGGGCTTCTCTCATC 360

QY 361 CGGAGAGCCAGACCAAGAGGCTTACTCTGTCAAGTCCGCTCAGCCGCCCTGCA 420
Db 361 CGGAGAGCCAGACCAAGAGGCTTACTCTGTCAAGTCCGCTCAGCCGCCCTGCA 420

QY 421 TCCTGGAGCCGATCAGACACTACAGATCACTGCTTGAACAATGGCTGGCTGTACATC 480

Db 421 TCCTGGAGCCGATCAGACACTACGATCCACTGCCTTGACAAATGGCTGGCTGATCATC 480
QY 481 TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGGTGACCACTTACTGTAGCTGGCG 540
Db 481 TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGGTGACCACTTACTGTAGCTGGCG 531
QY 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTGACAGGGGCTGGCCGCTCCCT 600
Db 532 -----GAGGGCTGGCCGCTCCCT 550
QY 601 GGCAAGGATATACCCCTACCTGTGACTGTGCAGAGACACCACTCAACTGAAAGAGCTG 660
Db 551 GGCAAGGATATACCCCTACCTGTGACTGTGCAGAGACACCACTCAACTGAAAGAGCTG 610
QY 661 GACAGCTCCCTCTGTTTCTGAGCTGCGCACAGGGAGAGTCTCTTCTCAGTGAAGGT 720
Db 611 GACAGCTCCCTCTGTTTCTGAGCTGCGCACAGGGAGAGTCTCTTCTCAGTGAAGGT 670
QY 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGA-CGAGGCTGTCTTTGATGA 779
Db 671 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAAGGAGGCTGTCTTTGATGA 730
QY 780 TGCCTAG 786
Db 731 TGCCTAG 737

RESULT 12
AF290986 737 bp mRNA linear PRI 21-JAN-2003
LOCUS AF290986 Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,
DEFINITION complete cds; alternatively spliced.
ACCESSION AF290986
VERSION AF290986.1 GI:17351922
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 737)
LORETO, M.P. and McGlade, C.J.
JOURNAL Cloning and characterization of human Src-like adaptor protein 2
MEDLINE and a novel splice isoform, SLAP-2-v
22415750 Oncogene 22 (2), 266-273 (2003)
PUBMED 12527895
REFERENCE 2 (bases 1 to 737)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
source 1..737
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1..633
/note="SLAP-2-v"
/codon_start=1
/product="Src-like adaptor protein-2 splice isoform"
/protein_id="AAL38198.1"
/db_xref="GI:17351923"
/translation="MGSLSRKRSLPSLSVSSVGGGPTMEARSKATAVALGSPF
AGGPAELSLRGEPLTIVSEDEGMMTVLSEVSGREYNIPSVHAKVSHGWLVEGLSRE
KAEELLLPKNPGAFILIRESGTRSGSYLSVRLSRPASWDRIRHYRIHCLDNGMLYI
SPRLTPSLQALVDHYSEGMPAPWQGYPTPTCDCAEDTQLERAGQLPVPF"

CDS
Query Match 84.4%; Score 663.4; DB 9; Length 737;
Best Local Similarity 93.4%; Pred. No. 1.1e-162;
ORIGIN

Matches 735; Conservative 0; Mismatches 1; Indels 51; Gaps 2;
QY 1 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTGTG 60
Db 1 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTGTG 60
QY 61 CAAGGCCAGGACCTGTGACCAATGGAAGCAGAGAGAGAAAGGCCACAGCCGTGGCCCTG 120
Db 61 CAAGGCCAGGACCTGTGACCAATGGAAGCAGAGAGAGAAAGGCCACAGCCGTGGCCCTG 120
QY 121 GGCAAGTTTCCCGCAGAGTGGCCCGCCGAGCTGTCCGTGAGACTCGGGAGCCATTGACC 180
Db 121 GGCAAGTTTCCCGCAGAGTGGCCCGCCGAGCTGTCCGTGAGACTCGGGAGCCATTGACC 180
QY 181 ATCGTCTTGAGATGAGACTGTGAGCGGTGTCTGAACTCAGGCAGAGAGATAT 240
Db 181 ATCGTCTTGAGATGAGACTGTGAGCGGTGTCTGAACTCAGGCAGAGAGATAT 240
QY 241 AACATCCCCAGCGTCCAGCTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGACC 300
Db 241 AACATCCCCAGCGTCCAGCTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGACC 300
QY 301 AGGGAAGAAAGCAGAGAACTGCTGTGTTACTCTGGGAACCTTGAGGGGCTTCTCATC 360
Db 301 AGGGAAGAAAGCAGAGAACTGCTGTGTTACTCTGGGAACCTTGAGGGGCTTCTCATC 360
QY 361 CGGAGAGCCAGACCAAGAGAGAGGCTTCTCTGATCCGCTCAGCCGCTGCA 420
Db 361 CGGAGAGCCAGACCAAGAGAGAGGCTTCTCTGATCCGCTCAGCCGCTGCA 420
QY 421 TCCTGGAGCCGATAGACACTACAGGATCCATGCTTGACAATGGCTGTGATCATC 480
Db 421 TCCTGGAGCCGATAGACACTACAGGATCCATGCTTGACAATGGCTGTGATCATC 480
QY 481 TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGGTGACCACTTACTCTGAGCTGGCG 540
Db 481 TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGGTGACCACTTACTCTGAGCTGGCG 531
QY 541 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCT 600
Db 532 -----GAGGGCTGGCCGCTCCCT 550
QY 601 GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGAAAGAGCTG 660
Db 551 GGCAAGGATATACCCCTACCTGTGACTGTGACAGAGAGACACCACTCAACTGAAAGAGCTG 610
QY 661 GACAGCTCCCTCTGTTTCTGAAAGTGGCCACAGGGAGAGAGTCTCTTCTCAGTGAAGGT 720
Db 611 GACAGCTCCCTCTGTTTCTGAAAGTGGCCACAGGGAGAGAGTCTCTTCTCAGTGAAGGT 670
QY 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGA-CGAGGCTGTCTTTGATGA 779
Db 671 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAAGGAGGCTGTCTTTGATGA 730
QY 780 TGCCTAG 786
Db 731 TGCCTAG 737

RESULT 13
BC052655 1384 bp mRNA linear ROD 04-NOV-2003
LOCUS BC052655 Mus musculus Src-like-adaptor 2, mRNA (cDNA clone MGC:60811
DEFINITION IMAGE:30040401), complete cds.
ACCESSION BC052655
VERSION BC052655.1 GI:30851667
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1384)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA),
Jonathan Keller (NCI, USA)
cDNA Library Preparation: Yulan Piao and Minoru Ko (National
Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Santripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAX Plate: 112 Row: e Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES
source

1. 1384
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="taxon:10090"
/clone="MGC:60811 IMAGE:30040401"
/tissue_type="Hematopoietic Stem Cell,
/lin-/c-Kit+/Sca-1-), mouse, 10 wks"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
/lin-/c-Kit+/Sca-1-) cDNA Library (Long)"
/lab_host="DH10B"
/note="Vector: pSPORT1"
1. 1384
/gene="Slaz2"
/note="synonyms: SLAP2, SLAP-2"

CDS
/db_xref="LocusID:77799"
/db_xref="MGI:1925049"
330. 1109
/gene="Slaz2"
/codon_start=1
/product="Slaz2 protein"
/protein_id="AAH52655.1"
/db_xref="GI:30851668"
/db_xref="LocusID:77799"
/db_xref="MGI:1925049"
/translation="MGSLSRGKTSPPSSSPDQEPVSMQPERKRYTAVAGSFP
GEQARLSRLGEPITTIISEDGWTVQSEVSGREYHMPVSVAKYAHGWLVEGLSREK
AEELILPENGAFILRESQTRGYSLSVRLSPASWDRIHYRIQLDNGWL
PRLTPSIHALVEHYSELADGICCPLEPCVLQKGLPGKDTPEPTVPTSLNWK
LDRSLFLEAPASGEASLSEGLRESLSYSISLAEDPLDA"

misc_feature

432. 596
/gene="Slaz2"
/note="SH3; Region: SH3 domain. SH3 (Src homology 3)
domains are often indicative of a protein involved in
signal transduction related to cytoskeletal organisation.
First described in the Src cytoplasmic tyrosine kinase.
The structure is a partly opened beta barrel"
/db_xref="CDD:pfam0018"
603. 899
/gene="Slaz2"
/note="SH2; Region: Src homology 2 domains"
/db_xref="CDD:cd00173"

ORIGIN

Query Match 68.6%; Score 539.4; DB 10; Length 1384;
Best Local Similarity 82.0%; Pred. No. 3.1e-130;
Matches 647; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

QY 1 ATGGGAAGTCTGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTGTC 60
Db 330 ATGGGAAGTCTGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTGTC 386
QY 61 CAAGGCCAGGACCTGTGACATGGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 387 CCAGACCAAGAACCCGTGTCTCAATGCAACAGAAAGAGAGAGAGAGAGAGAGAG 446
QY 121 GGCAATTTCCCGCAGGTGGCCCGCCAGCTGTGCTGAGACTGGGGAGCCATTGACC 180
Db 447 GGCAATTTCCCGCAGGTGGCCCGCCAGCTGTGCTGAGACTGGGGAGCCATTGACC 506
QY 181 ATCTCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 240
Db 507 ATCTCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 566
QY 241 AACATCCCCAGCGCTCCACGTGGCCCAAGTCTCCCATGGGTGGCTGATGAGGGCTGAGC 300
Db 567 CACATGCCAGTGTGTATGTGCTTAAGTCCGCCACGGGTGGCTGATGAGGGCTGAGC 626
QY 301 AGGAGAAAGCAGAGAACTGCTGTGTTAAGTGGAAACCTGAGAGGGCTTCCTCATC 360
Db 627 CGGAGAAAGCAGAGAACTGCTGTGTTAAGTGGAAACCTGAGAGGGCTTCCTCATC 686
QY 361 CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTCACTCCGCTCAGCCGCCCTGCA 420
Db 687 CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTCACTCCGCTCAGCCGCCCTGCA 746
QY 421 TCCTGGAGCCGATCAGACACTACAGATCCAGTGCCTTGACAAATGGCTGTACATC 480
Db 747 TCTTGGAGCCGATCAGACACTACAGATCCAGTGCCTTGACAAATGGCTGTACATC 806
QY 481 TCACCGCGCTCAGCTTCCCTCACTCCAGGCCCTGTGAGACCATTAAGTGTGAGCGG 540
Db 807 TCACCTGCGCTCAGCTTCCCTCACTCCAGGCCCTGTGAGACCATTAAGTGTGAGCGG 866
QY 541 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGCTGAGAGAGGCTGGCCGCTCCCT 600
Db 867 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGCTGAGAGAGGCTGGCCGCTCCCT 926

QY 601 GGCAAGATATATACCCCTACCTGTGTGACAGAGACCACTCAACTGGAAGAGCTG 660
Db 927 GGCAAGATATACACCTCCACCTGTGTGACCAACATCATCTAAATTGAAAAAGCTG 986
QY 661 GACAGCTCCCTCTCTGTTTCTGAAG--CTGCCACAGGGGAGAGTCTCTTCTCAGTGA 717
Db 987 GACCGAGCCTCTGTTTCTGGAAGCACTCGAGTGGGAGGACATCTCTGCTCAGTGA 1046
QY 718 GGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGAT 777
Db 1047 GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCTGGCTGAGGACCCCTTGATGATGCT 1106
QY 778 GATGCCTAG 786
Db 1107 TAGCCCTGG 1115
RESULT 14
AX511150 1348 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 1 from Patent WO0242452.
DEFINITION AX511150
ACCESSION AX511150
VERSION AX511150.1 GI:23392044
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Mcglade,J.C. and Loreto,M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 1 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source 1.1348
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 68.4%; Score 537.8; DB 6; Length 1348;
Best Local Similarity 81.9%; Pred. No. 8e-130;
Matches 646; Conservative 0; Mismatches 137; Indels 6; Gaps 2;
QY 1 ATGGGAAGTGTGCCAGCAGAGAAGAAATCTCTGCCAAGCCCAAGCTTCTCTGTC 60
Db 282 ATGGGAAGTGTGCCAGCAGAGAAGAAAC--CTCCAGCCCCAGCCCAAGCTCTCTGTC 338
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAGCAAGGCCACAGCCGTGGCCCTG 120
Db 339 CCAGACCAGGAACCCGTGTCTCATGCAACCAAGAAAGACACAAGGTACAGCTGTGGCCCTG 398
QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
Db 399 GGCAGTTTCCCGCAGGTGAACAGGCCAGACTATCTGTGAGACTCGGGGAGCCGCTGACC 458
QY 181 ATGCTCTTGAGGATGAGACTGTGACGCTGTGCTGAAGTCTCAGGACAGAGATAT 240
Db 459 ATCATCTTGAGGATGAGATTGTGACAGTCCAGTCGGAAGTCTCAGGACAGAGATAC 518
QY 241 AACATCCCGACGCTCCAGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC 300
Db 519 CACATGCCAGTGTATGTGGCTAAAGTCGCCACGCGGTGCTGTACGAGGCTGAGC 578
QY 301 AGGAGAAAGCAGAGAACTGCTGTTTACCTGGGAACCCCTGAGAGGGCCCTTCTCATC 360
Db 579 CGGAGAAAGCCGAGGAACTACTCTGTTTACCTGGGAACCCCGGAGGGGCTTCTCATC 638
QY 361 CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTGAGTCCGCTCAGCCGCCCTGCA 420
Db 639 CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTGAGTCCGCTCAGCCGCCCTGCA 698
QY 421 TCCTGGAGCCGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGCTGTACATC 480

Db 699 TCTTGGAGCCGATCAGACACTACAGGATACAGCGGTCTTGACAATGGCTGTACATC 758
QY 481 TCACCGGCTCACCCTTCCCTCACTCCAGGCCCTGTGGACCATTAATCTGAGCTGGCG 540
Db 759 TCACCTCGCTCACCCTTCCCTCACTCCAGGCCCTGTGGACCATTAATCTGAGCTGCA 818
QY 541 GATGACATCTGCTGCTTACTCAAGAGGCCCTGTGTCTGCAAGAGGGCTGGCCCTCCCT 600
Db 819 GATGCAATCTGCTGTCTCCCTCAAGGAGCCGTGTGTCTGCAAGAGCTTGGCCACTACT 878
QY 601 GGCAAGATATATACCCCTACCTGTGACTGTGACAGAGACCACTCAACTGGAAGAGCTG 660
Db 879 GGCAAGATATACCTCCACTGTGACTGTGACCAACATCATCTAAATTGAAAAAGCTG 938
QY 661 GACAGCTCCCTCTGTTTCTGAAG--CTGCCACAGGGGAGAGTCTCTTCTCAGTGA 717
Db 939 GACCGAGCCTCTGTTTCTGGAAGACACCTGCGAGTGGGAGGACATCTCTGCTCAGTGA 998
QY 718 GGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGAT 777
Db 999 GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCTGGCTGAGGACCCCTTGATGATGCT 1058
QY 778 GATGCCTAG 786
Db 1059 TAGCCCTGG 1067
RESULT 15
AF287467 1348 bp mRNA linear ROD 03-JUN-2002
LOCUS AF287467
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION AF287467.1 GI:17351918
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Loreto,M.P., Berry,D.M. and Mcglade,C.J.
TITLE Functional cooperation between c-Cbl and Src-like adaptor protein 2
JOURNAL Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
MEDLINE 22022020
PUBMED 12024036
REFERENCE 2 (bases 1 to 1348)
AUTHORS Loreto,M.P. and Mcglade,C.J.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
source 1.1348
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/dev_stage="day 15 embryo"
282..1061
/note="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AA138196.1"
/db_xref="GI:17351919"
/translation="MGSLSRSGKTSPPSSGDDQEPVSMQPERHKVTAVALGSFPA
GEQARLSRLGEPRLIISEDGDMWTQSEVSGREYHMPVYVAKVAHGMVYEGISREK
AEELLPLGNPGAFILRESQTRGYSLSVPSRPAWDRI RHYRIORLDNGMLYIS
PLTFPSLHALVEHYSELADGICCPLEBPVQLKGLPDPKDTPPVTVPTVYSSLNWKK
LDRSLFLLEAPASGBASLSEGLRESLSYSISLADPLDDA"
381..551
/note="Region: SH3 domain"
552..824
/note="Region: SH2 domain"
1301..1306
misc_feature
misc_feature
polya_signal

ORIGIN

Query Match 68.4%; Score 537.8; DB 10; Length 1348;
Best Local Similarity 81.9%; Pred. No. 8e-130;
Matches 646; Conservative 0; Mismatches 137; Indels 6; Gaps 2;

```
QY      1 ATGGGAGTCTGCCACAGAGAAATCTGCAAGCCCAAGCTTGAGTCTCTGTC 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      282 ATGGGAGTGTTCAGCAGAGGAAAC---CTCCAGCCCAAGCCCAAGCTCTGTGT 338

QY      61 CAAGGCCAGGACCTGTGACCATGAGACAGAGAAAGCAAGGCCACAGCCGTGCCCTG 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      339 CCAGACCAGGAACTGTCATGCAACCAAGAACACAAAGGTACAGCTGTGCCCTG 398

QY      121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      399 GGCAGTTTCCCGCAGGTGAACAGGCCAGACTATCTTGAACTCGGGAGCTGTAAC 458

QY      181 ATGCTCTGAGATGAGACTGTGTGACGGTGTGCTGAAGTCTCAGGCAAGATAT 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      459 ATCATCTGAGGATGAGATGTGTGACAGATCCAGTCCGAACTCAGGCAAGATAC 518

QY      241 AACATCCCCAGCGTCCAGTGCCCAAGTCTCCATGGGTGCTGTATGAGGCGCTGAGC 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      519 CACATGCCCAGTGTGTATGTGGCTAAAGTCGCCACGGGTGCTGTACGAGGCGCTGAGC 578

QY      301 AGGAGAAAGCAGAGAACTGCTGTGTTACTGTGGAACCTTGAGGGGCTTCTCATC 360
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      579 CCGGAGAAAGCCGAGGAATACTCTGTACTGGGAACCCCGAGGGGCTTCTCATC 638

QY      361 CCGGAGAGCCAGACCAAGGAGAGGCTTACTCTGTGATCCGCTCAGCCGCCCTGCA 420
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      639 CCGGAGAGCCAGACCAAGGAGGCTGTATTCCTGTCCGCTCCGACTCAGCCGCCCTGCA 698

QY      421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGTACATC 480
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      699 TCTTGGGACCGGATCAGACACTACAGGATCAGCGCTTGAACAATGGCTGTACATC 758

QY      481 TCAACCGCCTCAACCTTCCCTCACTCCAGGCCCCGTGTGACCACTTACTGTAGCTGGCG 540
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      759 TCACTCGCCTCAACCTTCCCTCACTCCAGCCTTGTGTGAGCACTTACTGTAGCTAGCA 818

QY      541 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTGTGAGAGGCGCTGCCCTCCCT 600
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      819 GATGACATCTGCTGCTACTCCCTCAAGGAGCCGTGTGTCTGTGAGCACTTACTGTAGCTAGCA 878

QY      601 GGCAGGATATACCCCTACCTGTGTGCTGTGACAGAGACACCACTCACTGGAAGAAGCTG 660
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      879 GGCAGGATATACCTTCACTGTGTGCTGTGACCAATCATCACTAAATGGAAAAAGCTG 938

QY      661 GACAGCTCCCTCTCTGTTTCTGAAG---CTGCCACAGGGGAGGAGTCTTCTCTAGTGAG 717
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      939 GACCGCAGCTCTCTGTTTCTGAAGCACTGCGAGTGGGGAGGCACTCTGTCTAGTGAG 998

QY      718 GGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGAT 777
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      999 GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCTGGCTGAGGAGCCCTTGATGATGCT 1058

QY      778 GATGCTTAG 786
      ||| ||| |||
Db      1059 TAGCCCTGG 1067
```

Search completed: November 16, 2004, 22:36:22
Job time : 3774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2004, 23:32:28 ; Search time 4546 Seconds

(without alignments)
2715.049 Million cell updates/sec

Title: US-10-043-649-2

Perfect score: 1351

Sequence: 1 MGSLSRKRSLPSPSLSSV.....RESLSFYISLNDEAVSLDDA 261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO_spool/US10043649/runat_16112004_060536_28719/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10043649 @CGN 1_1_3731 @runat_16112004_060536_28719 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1351	100.0	786	6	AX511153 Sequence
2	1351	100.0	786	6	AX572845 Sequence
3	1351	100.0	786	9	AF290985 Homo sapi
4	1351	100.0	786	9	AF326353 Homo sapi

5	1351	100.0	2415	9	AK025645 Homo sapi
6	1351	100.0	2538	9	BC042041 Homo sapi
7	1351	100.0	2567	6	AX452880 Sequence
8	1347	99.7	1183	6	AX443133 Sequence
9	1347	99.7	1183	6	AX443135 Sequence
10	1345	99.6	2788	6	AX780857 Sequence
11	1200.5	88.9	737	6	AX511155 Sequence
12	1200.5	88.9	737	6	AF290986 Homo sapi
13	1032	76.4	777	6	AX511151 Sequence
14	1032	76.4	1348	6	AX511150 Sequence
15	1032	76.4	1348	10	AF287467 Mus muscu
16	1032	76.4	1384	10	BC052655 Mus muscu
17	1023	75.7	1321	10	AF434990 Mus muscu
18	770	57.0	441	6	CQ724864 Sequence
19	647	47.9	864	6	CQ414231 Sequence
20	587	43.4	1571	5	BC056035 Xenopus l
21	505	37.4	1631	10	AY217759 Rattus no
22	496	36.7	921	10	MMU29056 Mus muscu
23	496	36.7	2648	10	BC032922 Mus muscu
24	496	36.7	2773	10	MMU131777 Mus muscu
25	495.5	36.7	841	10	AY079449 Mus muscu
26	495.5	36.7	841	10	AY079450 Mus muscu
27	491	36.3	1076	9	HSU30473 Homo sapien
28	491	36.3	1870	9	BC007042 Homo sapi
29	491	36.3	2021	9	HSU44403 Human Src-1
30	491	36.3	2109	6	AX428893 Sequence
31	491	36.3	2665	6	AX333017 Sequence
32	491	36.3	2665	9	D89077 Homo sapien
33	491	36.3	2670	6	CQ725810 Sequence
34	491	36.3	3090	9	HSM808530 Homo sapi
35	485.5	35.9	831	9	CR536537 Homo sapi
36	482	35.7	145068	9	HSU9777B1 Human DNA
37	482	35.7	145833	2	AC026539 Homo sapi
38	471.5	34.9	849	5	AY278230 Gallus ga
39	384.5	28.5	2703	10	MUSLYNB Rattus norv
40	379.5	28.1	1538	10	RATLYNBTYR Rattus norv
41	379.5	28.1	2302	10	RNAF000301 Rattus no
42	379.5	28.1	2824	10	RNAF000302 Rattus no
43	374.5	27.7	1361	6	CQ729540 Sequence
44	374.5	27.7	1539	12	AY335782 Synthetic
45	374.5	27.7	1613	9	BC075001 Homo sapi

ALIGNMENTS

RESULT 1	AX511153	786 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	AX511153	Sequence 4 from Patent WO0242452.			
DEFINITION	AX511153				
ACCESSION	AX511153				
VERSION	AX511153.1	GI:23392046			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	Mcglade, J.C. and Loreto, M.P.			
AUTHORS		Adapter gene			
TITLE		Patent: WO 0242452-A 4 30-MAY-2002;			
JOURNAL		The Hospital for Sick Children (CA)			
FEATURES					
source		Location/Qualifiers			
		1..786			
		/organism="Homo sapiens"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:9606"			

ORIGIN

Alignment Scores:
Pred. No.: 1.37e-101 Length: 786
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX511153 (1-786)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
DB 1 ATGGGAAGTCTGCCAGAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCTTGTCTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB 61 CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGCTGGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 121 GGCAGTTTCCCGGAGGTGGCCCGCCGAGCTGTCCGTGAGACTCGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
DB 181 ATCGTCTGTAGGATGGAGACTGTGTGACCGGTGCTGTCTGAAGTCTCAGGCAGAGATAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
DB 241 AACATCCCCAGCGCTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
DB 301 AGGGAAGAAACAGAGAACTGTGTGTACTCTGGGAACTTGAGGGGCTTCTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 361 CGGAGAGACCAGACCAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCGCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
DB 421 TCCTGGGACCGGATCAGACACTACGATCCACTGCCTTGACATGGCTGGCTGTACATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
DB 481 TCACCGCGCTCACTTCCCTCACTCCAGGCCCTGGTGACCATTACTTGAGCTGGCG 540
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB 541 GATGACATCTGTGCTTACTCAAGAGCCCTGTGTCTCTGACAGGGGCTGGCCCTCCCT 600
QY 201 GlnLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
DB 601 GGCAAGATATATACCTTACCTGTGACTGTGACAGAGACACCACTCACTGGAAGAAGCTG 660
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
DB 661 GACAGCTCCCTCCTCTTCTTGAAGCTGCCACAGGGAGAGAGTCTTCTCAGTGAGGCT 720
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
DB 721 CTCCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTGGATGAT 780
QY 261 Ala 261
DB 781 GCC 783

RESULT 2
AX572845 786 bp DNA linear PAT 29-NOV-2002
LOCUS AX572845
DEFINITION Sequence 1 from Patent WO02055707.
ACCESSION AX572845
VERSION AX572845.1 GI:26004935
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1

AUTHORS Holland, S.J., Mendenhall, M.K., Pardo, J., Spencer, C., Fu, A.C.,
Luo, Y., Payan, D.G., Mancebo, H.S., Wu, J., Zhou, X., Shen, M.,
Liao, X.C. and Sheng, N.
TITLE Cloning of an inhibitor of antigen-receptor signaling by a
retroviral-based functional screen
JOURNAL Patent: WO 02055707-A 1 18-JUL-2002;
Rigel Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..786
/note="unnamed protein product"
/translation="MGSLSRKRSLPSPSLSSVQGGPVTMEERKATATAVALGSPF
AGGPAELSLRGEPLTIVSEGDMMWTVLSEVSGREYNIPSVHAKVSHGWLVEGLSRE
KAEELILPNDGAFILRESQTRGSYLSVRSPASWDRIRHYRIHCLDNGWLYI
SPRLTFPSLQALVDHYSELADICCLKEPCVLRAPLPKGDIPLPVTVQRTPLNWK
ELDSSLFSFAATGEESLISEGLRESISFYISLNDKAVSLDDA"

CDS

/note="unnamed protein product"
/codon_start=1
/protein_id="CAD58542.1"
/db_xref="GI:26004936"
1..786
/translation="MGSLSRKRSLPSPSLSSVQGGPVTMEERKATATAVALGSPF
AGGPAELSLRGEPLTIVSEGDMMWTVLSEVSGREYNIPSVHAKVSHGWLVEGLSRE
KAEELILPNDGAFILRESQTRGSYLSVRSPASWDRIRHYRIHCLDNGWLYI
SPRLTFPSLQALVDHYSELADICCLKEPCVLRAPLPKGDIPLPVTVQRTPLNWK
ELDSSLFSFAATGEESLISEGLRESISFYISLNDKAVSLDDA"
Alignment Scores:
Pred. No.: 1.37e-101 Length: 786
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX572845 (1-786)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
DB 1 ATGGGAAGTCTGCCAGAGAAATAATCTGTGCCAAGCCCAAGCTTGAGTTCTTGTCTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB 61 CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGCTGGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 121 GGCAGTTTCCCGGAGGTGGCCCGCCGAGCTGTCCGTGAGACTCGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
DB 181 ATCGTCTGTAGGATGGAGACTGTGTGACCGGTGCTGTCTGAAGTCTCAGGCAGAGATAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
DB 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
DB 301 AGGGAAGAAACAGAGAACTGTGTGTACTCTGGGAACTTGAGGGGCTTCTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 361 CGGAGAGCCAGACCAGAGAGGCTTTTACTCTGTGACGTCCGCTCAGCCGCGCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
DB 421 TCCTGGGACCGGATCAGACACTACGATCCACTGCCTTGACAAATGGCTGGCTGTACATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
DB 481 TCACCGCGCTCACTTCCCTCACTCCAGGCCCTGGTGACCATTACTTGAGCTGGCG 540
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB 541 GATGACATCTGTGCTTACTCAAGAGAGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCCT 600

QY 201 GlyAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 601 GGCAAGATATACCCCTACTGTGCTGACAGAGACACCACTCACTGAAAGAGCTG 660
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 661 GACAGCTCCCTCCTGTTTCTTGAAGCTGCCACAGGAGAGTCTCTTCTCAGTGAGGCT 720
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTGTGATGAT 780
QY 261 Ala 261
Db 781 GCC 783
RESULT 3
LOCUS AF290985 786 bp mRNA linear PRI 21-JAN-2003
DEFINITION Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF290985
VERSION AF290985.1 GI:17351920
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Cloning and characterization of human Src-like adaptor protein 2
and a novel splice isoform, SLAP-2-v
JOURNAL Oncogene 22 (2), 266-273 (2003)
MEDLINE 22415750
PUBMED 12527895
REFERENCE 2 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
source
1. .786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1. .786
/note="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL38197.1"
/db_xref="GI:17351921"
/translation="MGSLPSRRKSLSPSLSSVQGGPVTMEERSKATAVALGSPF
AGGPAISLRLGEPITIVSEGDWMTVLSVSGREYNIPSVHAKVSHGMLYEGLSRE
KAEELLLPGNPGGAFILRESQTRGSYSLSVRLSRPASWDRIHRYRIHCLDNGWLYI
SPRLTFPSLQALVDHYSEIADIDICLLKEPCVLQAGPLPGKDIPLPTVQRTPLNWK
ELDSLLFSEATGEESSLSEGRSLSFYISLNDKAVSLDDA"
ORIGIN
Alignment Scores:
Pred. No.: 1.37e-101 Length: 786
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-043-649-2 (1-261) x AF290985 (1-786)
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGACAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCTGTC 60

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGACAGAGAGAAAGCAAGCCACAGCCGTGGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAGTTTCCCGCAGGTGCGCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
QY 1LeValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATGCTCTGTGAGATGAGACTGTGTGACGGTGTGTGAACTCTCAGGCAGAGAGTAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGlyLeuSer 100
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 301 AGGAGAAGACAGAGAACTGCTGTGTACCTGGGAACCTTGAGGGGCTTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGACCAAGAGAGGCTCTTACTCTGTGTGATCCGCTCAGCCGCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 421 TCCTGGAGCCGGATGACACACTACAGGATCCACTGCTTGAACAATGGCTGGCTGATCATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCTCACCTTCCCTCCTCAGGCTGTGTGACCATTAATTACTGTGAGCTGGCG 540
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 541 GATGACATCTGTGCTTCACTCAAGAGAGCCCTGTGTCTGCAGAGGCTGGCCGCTCCCT 600
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 601 GGCAAGATATACCCCTACTGCTGACTGTGCAGAGACACCACTCACTGGAAGAGACTG 660
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 661 GACAGCTCCCTCCTGTTTCTGAAAGCTGCCACAGGGAGAGACTCTTCTCACTGAGGCT 720
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTTGATGAT 780
QY 261 Ala 261
Db 781 GCC 783
RESULT 4
LOCUS AF326353 786 bp mRNA linear PRI 08-NOV-2001
DEFINITION Homo sapiens Src-like adapter protein-2 mRNA, complete cds.
ACCESSION AF326353
VERSION AF326353.1 GI:16797891
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Holland, S.J., Liao, X.C., Mendenhall, M.K., Zhou, X., Pardo, J.,
Chu, P., Spencer, C., Fu, A.C., Sheng, N., Yu, P., Pali, E., Nagin, A.,
Shen, M., Yu, S., Chan, E., Wu, X., Li, C., Woisetschlager, M.,
Aversa, G., Kolbinger, F., Bennett, M.K., Molineaux, S., Luo, Y.,
Payan, D.G., Mancebo, H.S.Y. and Wu, J.
TITLE Functional Cloning of Src-like Adapter Protein-2 (SLAP-2), a Novel
Inhibitor of Antigen Receptor Signaling
JOURNAL J. Exp. Med. 194 (9), 1263-1276 (2001)

MEDLINE 21553259
PUBMED 11696592
REFERENCE 2 (bases 1 to 786)
AUTHORS Holland,S.J., Mendenhall,M.K., Zhou,X., Spencer,C., Pardo,J., Fu,A.C., Sheng,N., Shen,M., Liao,C., Luo,Y., Payan,D.G., Mancebo,H.S.Y. and Wu,J.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand Avenue, South San Francisco, CA 94080, USA
FEATURES
source
1. .786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
1. .786
/note="SLAP2; SH2 and SH3 domain-containing adaptor"
/codon_start=1
/product="Src-like adapter protein-2"
/protein_id="AAL29204.1"
/db_xref="GI:16797892"
/translation="MGSLPSRRKSLPSPSLSSSVQGGPVTMEAEKSKATAVALGSPF AGPAELSLRLEPLTIVSEGDWMTVLSEVSGREYNIPSVHVAKVSHGWLIEGLSRE KAEELLLPGNPGAFILRESQTRGSYSLSVRLSRPASWDRIHRIHCLDNGWLYI SPRLTPSLQALVDHYSELADDICLLKEPCVLQRAGPLPKXDIPLPVTVQRTPLNWK ELDSSILFSEATGESESLSEGLRESLSFYISLUNDEAVSLDDA"
ORIGIN
Alignment Scores:
Pred. No.: 1.37e-101 Length: 786
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-043-649-2 (1-261) x AF326353 (1-786)
QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGCAGAAATAATCTCTGCCAAGCCCAAGCTTGAGTTCCCTGTC 60
QY 21 GlnGlyGlnGlyProValTherGluAlaGluArgSerLyAlaThrAlaValAlaLeu 40
Db 61 CAAGGCCAGGGAAGCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGlyTyr 80
Db 181 ATGTTCTTGAGGATGAGACTGGTGGACGGTGTCTGTGAGTCTCAAGGACAGAGATAT 240
QY 81 AsnIleProSerValHisValAlaValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCGCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCCTGAGC 300
QY 101 ArgGlyAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 301 AGGAGAAAGCAGAGAACTGCTGTGTACTCTGGAACCCCTGAGGGGCTTCTCATC 360
QY 121 ArgGlySerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGACAGAGAGAGGCTTACTCTGTCTAGTCCGCCCTCAGCCCGCTGCA 420
QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPheTyrIle 160
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCACTGCTTGAACAATGGCTGCTACATC 480
QY 161 SerProArgIleuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCCTCACTTCCCTCACTCCAGGCCCTGTGTGAGCAATTACTTGAAGCTGGCG 540

QY 181 AspAspIleCysCysLeuLeuLyGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 541 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTCTCTGCAGAGGCTGGCCGCTCCCT 600
QY 201 GlyLyAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLyGluLeu 220
Db 601 GGCAAGATATACCCCTAAGCTGTGACTGTGCAGAGACACCACTCACTGGAAGAGCTG 660
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGly 240
Db 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGAGTCTCTCTGAGTGAAGGT 720
QY 241 LeuArgGlySerLeuSerPheTyrIleSerLeuAsnAspGlyAlaValSerLeuAsp 260
Db 721 CTCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTCTTGATGAT 780
QY 261 Ala 261
Db 781 GCC 783
RESULT 5
AK025645 2415 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ21992 fis, clone HEP06554.
DEFINITION AK025645 AK025645.1 GI:10438227
ACCESSION AK025645
VERSION AK025645.1
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2415)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: f1cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES
source
1. .2415
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP06554"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="cloning vector pME18SFL3"
61. .846
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15201.1"
/db_xref="GI:10438228"
/translation="MGSLPSRRKSLPSPSLSSSVQGGPVTMEAEKSKATAVALGSPF AGPAELSLRLEPLTIVSEGDWMTVLSEVSGREYNIPSVHVAKVSHGWLIEGLSRE KAEELLLPGNPGAFILRESQTRGSYSLSVRLSRPASWDRIHRIHCLDNGWLYI SPRLTPSLQALVDHYSELADDICLLKEPCVLQRAGPLPKXDIPLPVTVQRTPLNWK

/gene="SLA2"
/codon_start=1
/product="Src-like-adaptor 2, isoform a"
/protein_id="AAH42041.1"
/db_xref="GI:27469843"
/db_xref="LocusID:84174"
/db_xref="MIM:606577"
/translation="MGSLSRRKSLPSPLSSSVQGGPVTMEERSKATAVALGSEF
AGPAELSLRGEPLTIVSEDDGMWTVLSEVSGREYNIPSVHVAKSHGWLXEGLSRE
KABELLLPGRGAFILIRISQTRGSLSVRLSRPASWDRIHRYRILHCLDNGWLYI
SPRLTFPSLQALVDHYSELADICCLKEPCVLQRAGPLPKDIPLVTVQRTPLNWK
ELDSLLFSEATGESLSLSEGLRESTLSFYISINDEAVSLDDA"

ORIGIN

Alignment Scores:
Pred. No.: 5.34e-101 Length: 2538
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-043-649-2 (1-261) x BC042041 (1-2538)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
Db 363 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 422
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 423 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGCCAGCCGTGGCCCTG 482
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 483 GGCAGTTTCCGGCAGGTGGCCGGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 542
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 543 ATCGTCTCTGAGATGAGACTGTGACCGTGTCTGTGAAGTCTCAAGCAGAGAGTAT 602
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 603 AACATCCCAAGCGTCCACGTGGCCAAAGTCTCCCATGAGTGGCTGTATGAGGGCCTGAGC 662
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 663 AGGGAGAAACAGAGAACTGCTGTGTTACTCTGGGAACCCCTGAGGGGCCCTTCCTCATC 722
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 723 CGGGAGGCCAGACCAAGAGAGGCTTACTCTCTGTCACTCCGCCCTCAGCCGCCCTGCA 782
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 783 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACAAATGGCTGGCTGTACATC 842
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 843 TCACCGCGGCTCACCTTCCCTCCTCAGGCCCTGTGGAGCCATTACTCTGAGCTGGCG 902
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 903 GATGACATCTGTGCTTACTCAAGAGAGCCCTGTGTCTGACAGAGGGCTGGCCCTCCCT 962
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 963 GGCAAGGATATACCCCTAAGTGTGACTGTGACAGAGACACCACTCACTGGAAGAAGCTG 1022
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 1023 GACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCTTCTCAGTGAAGGT 1082
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260

Db 1083 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGATGAT 1142
QY 261 Ala 261
Db 1143 GCC 1145

RESULT 7
AX452880
LOCUS AX452880 2567 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0242457.
ACCESSION AX452880
VERSION AX452880.1 GI:21712520
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Chang,H., Yang,W.P., Wu,Y., Whitney,G.S., Perez-Villar,J.J. and Kanner,S.B. Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific expression
JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;
Bristol-Myers Squibb Co. (US)
FEATURES
source location/Qualifiers
1..2567
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 5.41e-101 Length: 2567
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX452880 (1-2567)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
Db 415 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 474
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 475 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGCCAGCCGTGGCCCTG 534
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 535 GGCAGTTTCCGGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 594
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 595 ATCGTCTCTGAGATGAGACTGTGACCGGTGCTGTCTGAAGTCTCAGGACAGAGAGTAT 654
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 655 AACATCCCAAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGAGC 714
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 715 AGGAGAAAGCAGAGAACTGCTGTGTTACTTGGAACCCCTGAGGGGCCCTTCCTCATC 774
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 775 CGGGAGAGCCAGACCAAGAGGCTCTTACTCTGTCACTCCGCCCTCAGCCGCCCTGCA 834
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160

Db 835 TCCTGGACCGGATCAGACACTACGATCCACTGCCTTGACAATGGCTGGCTGACATC 894
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 895 TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGGTGACCACTACTCTGAGCTGGCG 954
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 955 GATGACATCTGCTGCTCACTCAAGAGCCCTGTGTCTGACAGGGCTGGCCCGCTCCCT 1014
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 1015 GGCAAGGATATACCCCTACCTGTGACTGTGCAGAGACACCACTCAACTGAAAGAGCTG 1074
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 1075 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTTCTTCACTGAGGCT 1134
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260
Db 1135 CTCCGGAGTCCCTCACTTCTCATCATCAGCCGTAATGACGAGCTGTCTCTTGATGAT 1194
QY 261 Ala 261
Db 1195 GCC 1197

RESULT 8
AX443133
LOCUS AX443133 1183 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 74 from Patent WO0216599.
ACCESSION AX443133
VERSION AX443133.1 GI:21690555
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.
Proteins and nucleic acids encoding same
Patent: WO 0216599-A 74 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)

TITLE
JOURNAL
Proteins and nucleic acids encoding same
Patent: WO 0216599-A 74 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)

FEATURES
source
1. .1183
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4.68e-101 Length: 1183
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.70% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX443133 (1-1183)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 398 ATGGGAAGTCTGCCACAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 457
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 458 CAAGGCCAGGAGCCTGTGACCATGGAAGACAGAGAAGCAAGGCCACAGCCGTGGCCCTG 517
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGAGCTGTGCTGAGACTCGGGAGCCATTGACC 577

QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 578 ATGTCCTCTGAGATGAGAGTGTGTGACGCGTGTCTGTAAGTCTCAGCAGAGAGTAT 637
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 638 AACATCCCCAGCGCTCCACGTGGGCAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 697
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 698 AGGAGAAGACAGAGGAACTGCTGTGTACCTGGGAACCTTGAGGGGCTTCTCATC 757
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 758 CGGAGAAGCCAGACCAGAGAGGCTTCTTACTCTGTGATCCGCCCTCAGCCGCTGCA 817
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 818 TCCTGGACCGGATCAGACACTACAGATCCACTGCTTGACAATGGCTGCTGTACATC 877
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 878 TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGGTGGAACCATTACTCTGAGCTGGC 937
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 938 GATGACATCTGCTGCTCACTCAAGAGACCCTGTGTCTGACAGAGGCTGGCCGCTCCCT 997
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 998 GGCAAGGATATACCCCTACCTGACTGTGCAGAGACACCACTCAACTGAAAGAGCTG 1057
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 1058 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGAGTCTTCTTCACTGAGGCT 1117
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260
Db 1118 CTCCGGAGTCCCTCACTTCTTACATCAGCTGAATGACGAGGCTGTCTCTTGATGAT 1177
QY 261 Ala 261
Db 1178 GCC 1180

RESULT 9
AX443135/c
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 76 from Patent WO0216599.
ACCESSION AX443135
VERSION AX443135.1 GI:21690556
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.
Proteins and nucleic acids encoding same
Patent: WO 0216599-A 76 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)

TITLE
JOURNAL
Proteins and nucleic acids encoding same
Patent: WO 0216599-A 76 28-FEB-2002;
Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)

FEATURES
source
1. .1183
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4.68e-101 Length: 1183
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0

Best Local Similarity: 99.62%
Query Match: 99.70%
DB: 6
Mismatches: 1
Indels: 0
Gaps: 0

US-10-043-649-2 (1-261) x AX443135 (1-1183)

QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal	20
Db	786	ATGGGAAGTCTGCCCAAGAAATCTTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	727
QY	21	GlInGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	726	CAAGGCCAGGAGCTGTGACCATGGAAGCAGAGAAAGCCACAGCCGTGGCCCTG	667
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
Db	666	GGCAGTTTCCCGGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC	607
QY	61	IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr	80
Db	606	ATCGTCTGTAGAGTGAAGACTGTGTGACGGTGTCTGAACTCTCAGGACAGAGATAT	547
QY	81	AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
Db	546	AACATCCCCAGCGCTCCACGTGGGCAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	487
QY	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuIle	120
Db	486	AGGAGAAAGCAGAGAACTGTGTGTACTTGGAACCTTGAGGGCCCTTCCTCATC	427
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	426	CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTCACTCCGCTCAGCCGCCCTGCA	367
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
Db	366	TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGAACAATGGCTGGCTGTATATC	307
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db	306	TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGGTGACCAATTACTCTGAGCTGGC	247
QY	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
Db	246	GATGACATCTGTGCTTACTCAAGAGCCCTGTGTCTGACAGGGCTGGCCGCTCCCT	187
QY	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db	186	GGCAAGATATACCCCTTACTGTGACTGTGACAGAGACACCACTCACTGGAAAGAGCTG	127
QY	221	AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGly	240
Db	126	GACAGTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGAGTCTTCTTCTCAGTGAAGGT	67
QY	241	LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp	260
Db	66	CTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTCTTGATGAT	7
QY	261	Ala 261	
Db	6	GCC 4	

RESULT 10
AX780857 2788 bp DNA linear PAT 14-JUL-2003

LOCUS
AX780857 Sequence 3014 from Patent WO03039443.

DEFINITION
AX780857

VERSION
AX780857.1 GT:32697851

KEYWORDS
SOURCE

ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
1
Haeflrich, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.
Novel genetic markers for leukemias
Patent: WO 03039443-A 3014 15-MAY-2003;
Deutsches Krebsforschungszentrum (DE) ;
Ludwig-Maximilian-Universitaet Muenchen (DE) ;
pd Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)

FEATURES
source
1. 2788
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.85e-100 Length: 2788
Score: 1345.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX780857 (1-2788)

QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal	20
Db	387	ATGGGAAGTCTGCCCAAGAAATCTTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	446
QY	21	GlInGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	447	CAAGGCCANGAGCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG	506
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
Db	507	GGCAGTTTCCCGGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC	566
QY	61	IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr	80
Db	567	ATCGTCTGTAGAGTGAAGACTGTGTGACGGTGTCTGTGAAGTCTCAGGACAGAGATAT	626
QY	81	AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
Db	627	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	686
QY	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuIle	120
Db	687	AGGAGAAAGCAGAGAGACTGTGTGTACTTGGAACCTTGAGGGGCCCTTCCTCATC	746
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	747	CGGAGAGCCCAAGACAGAGAGGCTTACTCTCTGTCACTCGGCTCAGCCGCCCTGCA	806
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
Db	807	TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACATGGCTGGCTGTACATC	866
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db	867	TCACCGCGCCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCAATTACTCTGAGCTGGC	926
QY	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
Db	927	GATGACATCTGTGCTTACTCAAGAGCCCTGTGTCTGACAGAGGGCTGGCCCTCCCT	986
QY	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db	987	GGCAAGATATACCCCTTACTGTGACTGTGACAGAGACACCACTCACTGAAAGAGCTG	1046
QY	221	AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGly	240
Db	1047	GACAGTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGAGAGTCTCTTCTCAGTGAAGGT	1106

QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspasp 260
Db 1107 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGCTCTTGATGAT 1166
QY 261 Ala 261
Db 1167 GCC 1169

RESULT 11
AX511155
LOCUS AX511155 737 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 6 from Patent WO0242452.
ACCESSION AX511155
VERSION AX511155.1 GI:23392047
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS McGlade, J.C. and Loreto, M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 6 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source location/Qualifiers
1..737
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2.71e-89 Length: 737
Score: 1200.50 Matches: 241
Percent Similarity: 92.40% Conservative: 2
Best Local Similarity: 91.63% Mismatches: 1
Query Match: 88.86% Indels: 19
DB: 6 Gaps: 1

US-10-043-649-2 (1-261) x AX511155 (1-737)

QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCCGACAGAAATACTCTGCCAAGCTTGAGTCTCTGTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAGTTTCCCGCGAGGTGGCCCGCGAGCTGCTGAGACTCGGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTyrTyrThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCGTCTGTAGAGATGAGACTGTGTGACCGTGTCTGAGTCTCAGGCGAGAGATAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCCCAGGCTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 301 AGGAGAAAGCAGAGAACTGCTGTGTGTTACTGGGAAACCTGAGGGGCTTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGACCAGAGAGGCTTTACTTCTGTCAAGTCCGCTCAGCCGCCCTGCA 420
QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 421 TCCTGGGAGCCGATCAGACACTACAGAGATCCACTGCTTGACATGCTGCTGATCATC 480

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGGCTCACCCTTCCCTCACTCCAGGCCCTGTGTGACCATTA----- 528
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200
Db 529 -----TCGAGGGCTGCGCCGCTCCC 549
QY 200 OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLe 220
Db 550 TGGCAAGGATATACCCCTACCTGTGACTGTGCGGAGGACACCACTCAACTGGAAGAGCT 609
QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluG 240
Db 610 GGACAGCTCCCTCCTGTTTCTGAAAGCTGCCACAGGAGGAGACTCTTCTCAGTGAAGG 669
QY 240 YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAsp 260
Db 670 TCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGCGAGGCTGTCTCTTGATG 729
QY 260 spAla 261
Db 730 ATGCC 734

RESULT 12
AF290986
LOCUS AF290986 737 bp mRNA linear PRI 21-JAN-2003
DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,
complete cds; alternatively spliced.
ACCESSION AF290986
VERSION AF290986.1 GI:17351922
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 737)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Cloning and characterization of human Src-like adaptor protein 2
and a novel splice isoform, SLAP-2-v

JOURNAL Oncogene 22 (2), 266-273 (2003)
MEDLINE 22415750
PUBMED 12527895
REFERENCE 2 (bases 1 to 737)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

FEATURES
source location/Qualifiers
1..737
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1..633
/note="SLAP-2-v"
/codon_start=1
/product="Src-like adaptor protein-2 splice isoform"
/protein_id="AAL38198.1"
/db_xref="GI:17351923"
/translation="MGSLPSRRKSLPSLSVVOGGPVTWAEERSKATAVALGSFP
AGGPABLRLGEPLTIVSDGDMWTVLSEVSGREYNIPSVHYAKVSHGWLVEGLSRE
KAEILLILPGNPGAFILRESQTRGSYSLSVRLSRPASWDRIIRHYRIHCLDNGWLYI
SPRLTFPSLQALVDHYSEGWPAPWQGYTPTCDCAEDTTQLERAQLPVPF"

CDS

QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCCGACAGAAATACTCTGCCAAGCTTGAGTCTCTGTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAGTTTCCCGCGAGGTGGCCCGCGAGCTGCTGAGACTCGGGGAGCCATTGACC 180
QY 61 IleValSerGluAspGlyAspTyrTyrThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCGTCTGTAGAGATGAGACTGTGTGACCGTGTCTGAGTCTCAGGCGAGAGATAT 240
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCCCAGGCTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 301 AGGAGAAAGCAGAGAACTGCTGTGTGTTACTGGGAAACCTGAGGGGCTTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGACCAGAGAGGCTTTACTTCTGTCAAGTCCGCTCAGCCGCCCTGCA 420
QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 421 TCCTGGGAGCCGATCAGACACTACAGAGATCCACTGCTTGACATGCTGCTGATCATC 480

ORIGIN

Alignment Scores:
Pred. No.: 2.71e-89 Length: 737
Score: 1200.50 Matches: 241
Percent Similarity: 92.40% Conservative: 2
Best Local Similarity: 91.63% Mismatches: 1
Query Match: 88.86% Indels: 19

DB:	9	Gaps:	1
US-10-043-649-2 (1-261) x AF290986 (1-737)			
QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	1	ATGGGAAGTCTGCCACGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTG	60
QY	21	GIinglYinglYProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	61	CAAGGCCAGGACCTGTGATGCCATGGAGACAGAGAGCAAGGCCACAGCCGTGCCCTG	120
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
Db	121	GGCAGTTTCCCGGAGGTGGCCCGGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC	180
QY	61	IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr	80
Db	181	ATCGTCTGAGGATGAGACTGTGTGACGGTGTCTGTGAAGTCTCAGGCCAGAGATAT	240
QY	81	AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
Db	241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	300
QY	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle	120
Db	301	AGGAGAAAGCAGAGAACTGTGTGTACTCTGGGAACCTTGAGGGCCTTCTCATTC	360
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	361	CGGAGAGCCAGACCAAGAGAGGCTTACTCTGTGATCCGCCCTCAGCCGCCCTGCA	420
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
Db	421	TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACAAATGGCTGTACATC	480
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db	481	TCACCGCGCCTCACCTTCCCTCTCAAGGCCCTGTGGACCATTTAC-----	528
QY	181	AspAspIleCysCysLeuLeuGluProCysValLeuGln-ArgAlaGlyProLeuPr	200
Db	529	-----TGTGAGGGCTGGCCCGCTCCC	549
QY	200	OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe	220
Db	550	TGGCAAGATATACCCCTACCTGTGACTGTGCGGAGACACCACTCACTGAAAGAGCT	609
QY	220	uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluG	240
Db	610	GGACAGCTCCCTCTCTGTTTCTGAGCTGCCACAGGGAGAGTCTCTCTCAGTGAGGG	669
QY	240	YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA	260
Db	670	TCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGAGCGAGGCTGTCTTTGATG	729
QY	260	sPAla 261	
Db	730	ATGCC 734	
RESULT 13			
LOCUS	AX511151	777 bp	DNA linear PAT 27-SEP-2002
DEFINITION	Sequence 2 from Patent WO0242452.		
ACCESSION	AX511151		
VERSION	AX511151.1	GI:23392045	
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Mus musculus		
REFERENCE	1 Mcglade, J.C. and Loreto, M.P.		
AUTHORS			

TITLE		Adapter gene	
JOURNAL		Patent: WO 0242452-A 2 30-MAY-2002;	
FEATURES		The Hospital for Sick Children (CA)	
source		Location/Qualifiers	
		1..777	
		/organism="Mus musculus"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:10090"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1.83e-75	Length:	777
Score:	1032.00	Matches:	209
Percent Similarity:	85.88%	Conservative:	16
Best Local Similarity:	79.77%	Mismatches:	33
Query Match:	76.39%	Indels:	4
DB:	6	Gaps:	3
US-10-043-649-2 (1-261) x AX511151 (1-777)			
QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	1	ATGGGAAGTTGTCCAGCAGAGGAAAACC--TCCAGCCCCAGCCCCAGCTCCTGTGT	57
QY	21	GInGlyGInGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	58	CCAGACCAGGAACCCGTGTCCATGCACAACGAAAGACACAAGGTCACAGCTGTGGCCTG	117
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
Db	118	GGCAGTTTCCAGCAGGTGAACAGCCAGACTATCTTGAGACTCGGGAGCCGTGACC	177
QY	61	IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr	80
Db	178	ATCATCTCTGAGATGAGATTGGTGGACAGTCCAGTGGGAAGTCTCAGGCAGAGAGTAC	237
QY	81	AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
Db	238	CACATGCCAGTGTGTATGTGGCTAAAGTCGCCACGGGTGGCTGTACGAGGGCCTGAGC	297
QY	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle	120
Db	298	CGGAGAAAGCCGAGGAATCTCTGTACCTGGGAACCCCGGAGGCTTCTCTCATC	357
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	358	CGGAGAGCCAGACCAAGAGAGGCTGCTATTCCCTGTCCGCTCCGACTCAGCCGCCCTGCA	417
QY	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
Db	418	TCTTGGACCGGATCAGACACTACAGGATACAGCGTCTTGACAAATGGCTGTACATC	477
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db	478	TCACCTCGCCTCACCTTCCCTCACTCCACGCCCTTGGTGAGCATTACTCTGAGCTAGCA	537
QY	181	AspAspIleCysCysLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro	200
Db	538	GATGGCATCTGCTGTCCCTCAGGAGCCGTGTCTCTGAGAGGCTTGGGCCACTACCT	597
QY	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db	598	GGCAAAAGATACACCTCCACCTGTGACTGTGCCAACATCATCACTAAATTGAAAAAGCTG	657
QY	221	AspSerSerLeuLeuPheSerGluAla--AlaThrGlyGluGluSerLeuSerGlu	239
Db	658	GACCGAGCCCTCCTGTTTCTGGAAGCACTCGAGTGGGAGGCATCTCTGCTCAGTGAG	717
QY	240	GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp	259
Db	718	GGGCTCCAGAGTCCCTCAGTTCCTACATCAGCCTGGCTGAGGAC-----CCCTTGAT	771
QY	260	AspAla 261	

Db 772 GATGCT 777

RESULT 14
LOCUS AX511150 1348 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1 from Patent WO0242452.
ACCESSION AX511150
VERSION AX511150.1 GI:23392044
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS McGlade, J.C. and Loreto, M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 1 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source location/Qualifiers
1..1348
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 3.47e-75 length: 1348
Score: 1032.00 Matches: 209
Percent Similarity: 85.88% Conservative: 16
Best Local Similarity: 79.77% Mismatches: 33
Query Match: 76.39% Indels: 4
DB: 6 Gaps: 3

US-10-043-649-2 (1-261) x AX511150 (1-1348)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 282 ATGGGAAGTTGTTCACAGAGGAAACC--TCCAGCCCCAGCCCGCTCTGTGT 338

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 339 CCAGACCAGGAACCCGTGTTCATGCAACCAAGAACACAAAGTCAACAGCTGTGCCCTG 398

QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 399 GGCAAGTTCCAGCAGGTGAACAGGCCAGACTATCTCTGAGACTCGGGGAGCCGCTGACC 458

QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 459 ATCATCTCTGAGGATGAGATTGTGTGACAGTCCAGTCCGAAAGTTCAGGACAGAGATAC 518

QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 519 CACATGCCCAAGTGTATGTGGCTAAAGTCGCCCAAGGCTGTGACGAGGCTGAGC 578

QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 579 CGGAGAAAGCCGAGGAAGTACTCTGTTACTCTGGGAACCCCGAGGGGCTCTCTCATC 638

QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 639 CGGAGAGCCAGACAGAGAGAGCTGCTATTCCTGCTCCGACTCAGCCGCTGCA 698

QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 699 TCTTGGAGCCGATCAGACACTACAGATACAGCGCTCTTGACAAATGGCTGTACATC 758

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlyLeuAla 180
Db 759 TCACCTCGCCTCACTTCCCTCCTCCTCCTGAGCATTACTCTGAGCTAGCA 818

QY 181 AspAspIleCysCysLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200

Db 819 GATGGCATCTGCTGTCCCTCAGGAGACCGCTGTGTCTGCAGAAAGCTTGGCCACTACT 878

QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 879 GGCAAGATACACCTCCACTCTGACTGTGCCAACATCATCAATAATTGAAAAAGCTG 938

QY 221 AspSerSerLeuLeuPheSerGluAla--AlaThrGlyGluGluSerLeuSerGlu 239
Db 939 GACCGCAGCCTCCTGTTTCTGGAAGCACCTGCGAGTGGGAGGCATCTGCTCAGTAGAG 998

QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 999 GGGCTCCGAGAGTCCCTCAGTTCTCATCATCAGCCCTGGCTGAGGAC-----CCCTTGAT 1052

QY 260 AspAla 261
Db 1053 GATGCT 1058

RESULT 15
LOCUS AF287467 1348 bp mRNA linear ROD 03-JUN-2002
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION AF287467.1 GI:17351918
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1348)
JOURNAL Functional cooperation between c-Cbl and Src-like adaptor protein 2
MEDLINE Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
PUBMED 22022020
2 (bases 1 to 1348)
Loreto, M.P. and McGlade, C.J.
Direct Submission
Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

FEATURES
source location/Qualifiers
1..1348
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/dev_stage="day 15 embryo"
282..1061
/note="SIAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL38196.1"
/db_xref="GI:17351919"
/translation="MGSLSSRGKTSSPSSSGPDQEPVSMQPERHKVTAVAGSFPA
GEQARLSRLGEPLTIISEDDMWTVQSEVSGREYHMPVYVAKVAHGWLYEGLSREK
AEELLPLGNPGAFILRESQTRGCSLSVRLSRPASWDRIRHYRIQRLDNGWLYIS
PLTFPSLHALVEHYSELADGICCPLRPCVLQKLGPLPGKDTPTTPVTPTSSLNWK
LDRSLFLFEPASGEASLSEGLRESLSSYISLAEDPLDDA"
381..551
/note="Region: SH3 domain"
552..824
/note="Region: SH2 domain"
1301..1306

ORIGIN

polyA_signal

misc_feature
misc_feature
misc_feature

Alignment Scores:
Pred. No.: 3.47e-75 length: 1348
Score: 1032.00 Matches: 209
Percent Similarity: 85.88% Conservative: 16
Best Local Similarity: 79.77% Mismatches: 33
Query Match: 76.39% Indels: 4
DB: 10 Gaps: 3

US-10-043-649-2 (1-261) X AF287467 (1-1348)

QY		1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db		282	ATGGGAAGTTGTTCACAGAGAGGAAACC--TCCAGCCCCAGCCCCAGCTCCTTGCT	338
QY		21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db		339	CCAGACCAGAACCCCGTGTCCATGCACAACCAAGAACAACAAGGTCAAGCTGTGCCCTG	398
QY		41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
Db		399	GGCAGTTTCCCAGCAGGTGAACAAGGCCAGACTATCTCTGAGACTCGGGGAGCCGCTGACC	458
QY		61	IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr	80
Db		459	ATCATCTCTGAGGATGAGATTGTGTGACAGCTCCAGTCCGAAGTCTCAGGCACAGAGTAC	518
QY		81	AsnIleProSerValHisValAlaLysValSerHisGlyTyrPleuTyrgluGlyLeuSer	100
Db		519	CACATGCCCCAGTGTGTATGTGGCTAAGTCGCCACCGGGTGTGTACGAGGGCCCTGAGC	578
QY		101	ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle	120
Db		579	CGGAGAAAAGCCGAGGAATACTCCTGTACCTGGGAACCCGAGAGGGCCCTTCCTCATC	638
QY		121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db		639	CGGAGAGCCAGACCAGAGAGAGGCTGTATTCCCTGTCCGTCCGACTCAGCCGCCCTGCA	698
QY		141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
Db		699	TCTTGGGACCGGATCAGACACTACAGATACAGCGTCTTGACAATGGCTGGCTGTACATC	758
QY		161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db		759	TCACTTGCCCTCACCTTCCCTCACTCCACGCGCTTGGTGAAGCATTA CTCTGAGCTAGCA	818
QY		181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
Db		819	GATGGCATCTGTGCTGCCCTCAGGAGCCGTGTGTCTCTGACAGAAAGCTTGGGCCACTACT	878
QY		201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db		879	GGCAAAGATACACCTCCACCTGTGACTGTGCCAACATCATCACTAAATTGAAAAAGCTG	938
QY		221	AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuLeuSerGlu	239
Db		939	GACCGCAGCCTCCTGTTCTTGAAGACACCTGCGAGTGGGGAGGCATCTCTGCTCAGTAG	998
QY		240	GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuasp	259
Db		999	GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCTGGCTGAGGAC-----CCCTTGAT	1052
QY		260	AspAla 261	
Db		1053	GATGCT 1058	

```
Search completed: November 17, 2004, 02:15:10
Job time : 4555 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2004, 00:38:19 ; Search time 3371 seconds

(without alignments)
2821.350 Million cell updates/sec

Title: US-10-043-649-2
Perfect score: 1351
Sequence: 1 MGSLPSRRKSLPSPSSSV.....RESLSFYISLNDEAVSLDDA 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-Q=/cg2_1/USPTO_spool/US10043649/runat_16112004_060536_28728/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10043649 @CGN 1 1 3437 @runat_16112004_060536_28728 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1277	94.5	1002	5	BQ052308	BQ052308 AGENCOURT
2	1187	87.9	1069	5	BQ052468	BQ052468 AGENCOURT
3	1152.5	85.3	1020	5	BQ054281	BQ054281 AGENCOURT
4	1032	76.4	2637	3	AK088672	AK088672 Mus muscu
5	1028	76.1	2974	3	AK030877	AK030877 Mus muscu
6	937.5	69.4	926	3	AK020837	AK020837 Mus muscu
7	882.5	65.3	660	6	BY742155	BY742155 BY742155
8	871	64.5	1032	1	AL541041	AL541041 AL541041
9	711	52.6	566	4	BG284179	BG284179 602408226

10	684.5	50.7	986	5	BQ054265	BQ054265 AGENCOURT
11	611.5	45.3	960	5	BU944126	BU944126 AGENCOURT
12	564.5	41.8	878	5	BQ053486	BQ053486 AGENCOURT
13	561	41.5	377	1	AA959151	AA959151 v251906.r
14	560.5	41.5	660	2	BB635615	BB635615 BB635615
15	555	41.1	597	1	AL844311	AL844311 AL844311
16	550	40.3	778	4	BG178487	BG178487 602328305
17	545	40.2	606	1	AL844309	AL844309 AL844309
18	543	39.4	614	1	AL844307	AL844307 AL844307
19	532	39.3	781	7	CK596391	CK596391 AGENCOURT
20	531	39.3	569	6	CB426333	CB426333 601508 MA
21	521.5	38.6	627	2	BB619854	BB619854 BB619854
22	520	38.5	762	5	BP162888	BP162888 BP162888
23	518	38.3	701	7	CK833360	CK833360 4057315 B
24	496	36.7	2810	3	AK036167	AK036167 Mus muscu
25	495.5	36.7	831	9	AY416279	AY416279 Mus muscu
26	495	36.6	1386	3	AK041565	AK041565 Mus muscu
27	493.5	36.5	831	9	AY416277	AY416277 Homo sapi
28	491	36.3	1762	3	CR617843	CR617843 full-length
29	482	35.7	1037	1	AL539427	AL539427 AL539427
30	481.5	35.6	655	7	CF913437	CF913437 A0648D04-
31	480.5	35.6	1997	3	AK037901	AK037901 Mus muscu
32	477	35.3	966	1	AL551370	AL551370 AL551370
33	474.5	35.1	701	5	BX849096	BX849096 BX849096
34	472	34.9	1015	5	BX436423	BX436423 BX436423
35	471	34.9	972	5	BQ707614	BQ707614 AGENCOURT
36	468	34.6	1054	1	AL549826	AL549826 AL549826
37	461	34.1	656	7	CN792083	CN792083 4126904 B
38	460	34.0	796	7	CO574805	CO574805 AGENCOURT
39	456.5	33.8	871	5	BQ436143	BQ436143 AGENCOURT
40	451.5	33.4	775	6	CB938697	CB938697 IPCCJx13
41	448.5	33.2	934	7	CK411626	CK411626 AUF_IphdK
42	446	33.0	1063	5	BQ072745	BQ072745 AGENCOURT
43	445	32.9	723	5	BM950089	BM950089 UI-M-EHOP
44	443.5	32.8	777	6	CB938723	CB938723 IPCCJx13
45	443.5	32.8	882	5	BX415149	BX415149 BX415149

ALIGNMENTS

RESULT 1
BQ052308
LOCUS BQ052308 1002 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT_6868571 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5933542
5', mRNA sequence.
ACCESSION BQ052308 GI:19811648
VERSION BQ052308.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1002)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLM2118 row: d column: 23
High quality sequence stop: 670.
Location/Qualifiers
1. 1002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933542"

FEATURES

source
1. 1002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933542"

/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	1.89e-116	Length:	1002
Score:	1277.00	Matches:	257
Percent Similarity:	97.74%	Conservative:	2
Best Local Similarity:	96.98%	Mismatches:	2
Query Match:	94.52%	Indels:	4
DB:	5	Gaps:	0

US-10-043-649-2 (1-261) x BQ052308 (1-1002)

QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	96	ATGGGAAGTGTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC	155
QY	21	GlNGlYGlNGlYProValThrMetGlUAlaGlUArgSerLysAlaThrAlaValAlaLeu	40
Db	156	CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAAAGCAAGCCAGCCGTGCCCTG	215
QY	41	GlySerPheProAlaGlyGlyProAlaGlUleuSerLeuArgLeuGlYGlUProLeuThr	60
Db	216	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC	275
QY	61	IleValSerGlUAspGlyAspTrpTrpThrValLeuSerGlUValSerGlyArgGluTyr	80
Db	276	ATCGTCTGTAGATGAGACTGGTGGACGGTGCTGTGAGTCTCAAGCAGAGAGTAT	335
QY	81	AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGlUglyLeuSer	100
Db	336	AACATCCCAAGCTCCACGTGGCCAAATCTCCATGGGTGTATGAGGGCCTGAGC	395
QY	101	ArgGlUAlaGlUleuLeuLeuProGlYAsnProGlYAlaPheLeuIle	120
Db	396	AGGGAAGAAGCAGAGAACTGTGTGTAACCTGGAAACCTGGAGGGCCTTCTCATC	455
QY	121	ArgGlUArgGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	456	CGGGAGAGCCAGACCAGAGAGGCTCTTACTCTGTCAAGTCCGCTCAGCCGCTGCA	515
QY	141	SerTyrPaspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
Db	516	TCCTGGAGCCGATCAGACACTACAGATCTACTGCTTGACAAATGGCTGGCTGATCATC	575
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db	576	TCACCGCGCTCACTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTTGAGCTGGCG	635
QY	181	AspAspIleCysCysLeuLeuLysGlUProCysValLeuGlnArgAlaGlyProLeuPro	200
Db	636	GATGACATCTGCTGCTCACTCAAGAGAGCCCTGTGTCTGCAGAGGGGCTGGCCCTCCCT	695
QY	201	GlYlYsAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db	696	GGCAAGATATACCTCACTGTGACTGTGCAAGAGACACCACTCAACTGGAAGAAGCTG	755
QY	221	AspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGlUglUSerLeuSerGlu-G	240
Db	756	GACAGCTCCCTCTCTTTTCTGAAGCTGCCACAGGGGAGAGTCTCTCACTGAGGG	815
QY	240	lYleuArg-GlUserLeuSerPheTyrIleSerLeuAsnAspGlu-AlaValSerLeuAs	259

Db	816	GTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCCTGATGACGAGGCTGTCTTTGA	875
QY	259	pAspAla	261
Db	876	TGATGCC	882

RESULT 2
BQ052468
LOCUS
DEFINITION
AGENCOURT_6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
5', mRNA sequence.
BQ052468
BQ052468.1 GI:19811808
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1069)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCM218 row: n column: 13
High quality sequence stop: 681.
Location/Qualifiers
1. 1069
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933772"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

FEATURES

source

Alignment Scores:
Pred. No.: 1.93e-107 Length: 1069
Score: 1187.00 Matches: 244
Percent Similarity: 92.42% Conservative: 0
Best Local Similarity: 92.42% Mismatches: 1
Query Match: 87.86% Indels: 20
DB: 5 Gaps: 1

ORIGIN

US-10-043-649-2 (1-261) x BQ052468 (1-1069)

QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	89	ATGGGAAGTGTGCCAGCAGAAATAATCTTGCCCAAGCCCAAGCTTGAGTCTCTGTC	148
QY	21	GlNGlYGlNGlYProValThrMetGlUAlaGlUArgSerLysAlaThrAlaValAlaLeu	40
Db	149	CAAGGCCAGGAGCCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGCCCTG	208
QY	41	GlySerPheProAlaGlyGlyProAlaGlUleuSerLeuArgLeuGlYGlUProLeuThr	60
Db	209	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC	268

QY 61 ILevalSerGluAspGlyAspTrpThrValleuSerGluValSerGlyArgGluTyr 80
Db 269 ATCGTCTCTGAGGATGAGACTGTGTGACGGTGTCTGTGAAGTCTCAGGCGACAGAGTAT 328
QY 81 AsnIleProSerValHisValAlaIalysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 329 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCCTGAGC 388
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 389 AGGAGAAACAGACAGAACTGTGTGTTCCTGGAAACCTGGAGGGCCTTCTCATC 448
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 449 CGGAGAGCCAGACAGAGAGGCTTACTCTGTCTGAGTCCGCTCAGCCGCTGCA 508
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 509 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACAAATGGCTGGCTGTACATC 568
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeu-ValAspHisTyrSerGluLeuAl 180
Db 569 TCACCGCGCTCACCCTTCCCTCACTCCAGGCCCTGGGTGGACCACTTACTCTGAGG-- 625
QY 180 aAspAspIleCysCysLeuLeuLysGluProCysValleuGlnArgAlaGlyProLeuPr 200
Db 626 -----GCTGGCCCCGCTCCC 639
QY 200 cGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220
Db 640 TGGCAAGATATACCCCTACTGTGACTGTGACAGAGACACCACTCAACTGGAAGAGCT 699
QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluGluSerLeuLeuSerGluG 240
Db 700 GGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGAGAGTCTTCTTCTGAGG 759
QY 240 LyleuArgGlu-SerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 760 GTCTCCGGGAGAGTCCCTGACTTCTACATCAGCTGAATGACGAGGCTGTCTTTTGAT 819
QY 260 AspAla 261
Db 820 GATGCC 825
RESULT 3
BQ054281 1020 bp mRNA linear EST 29-MAR-2002
LOCUS BQ054281
DEFINITION AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
5', mRNA sequence.
ACCESSION BQ054281
VERSION BQ054281.1 GI:19813621
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1020)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2125 row: J column: 11
High quality sequence stop: 556.
FEATURES
Location/Qualifiers

source 1. .1020
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

Alignment Scores:
Pred. No.: 4.93e-104 Length: 1020
Score: 1152.50 Matches: 238
Percent Similarity: 93.51% Conservative: 7
Best Local Similarity: 90.84% Mismatches: 10
Query Match: 85.31% Indels: 7
DB: 5 Gaps: 1
US-10-043-649-2 (1-261) x BQ054281 (1-1020)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal 20
Db 212 ATGGGAAGTCTGCCAGCAGAAATAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 271
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 272 CAAGGCCAGGAGCTGTGACCATGGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTG 331
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 332 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGGCCATGACC 391
QY 61 ILevalSerGluAspGlyAspTrpThrValleuSerGluValSerGlyArgGluTyr 80
Db 392 ATCGTCTCTGAGATGAGACTGTGTGACCGGTGTCTGAAGTCTCAGGCGAGAGTAT 451
QY 81 AsnIleProSerValHisValAlaIalysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 452 AACATCCCCAGCGTCCACAGTGGCCAAAGTCTCCCATGGCTGTATGAGGGCCTGAGC 511
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 512 AGGAGAAAGCAGAGAACTGTGTGTTCCTGAGTCCAGAGGGGCTTCTCATC 571
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 572 CGGAGAGCCAGACCAAGAGAGGCTTACTCTGTCTGAGTCCGCTCAGCCGCTGCA 631
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 632 TCCTGNGACCGGATCAGACACTACAGATCCACTGCTTGACAAATGGCTGTACATC 691
QY 161 SerProArgLeuThrPheProSerLeuGlnAla-LeuValAspHisTyrSerGluLeuAl 180
Db 692 TCACCGCGCTCACCCTTCCCTCACTCCAGGCCCTGTGTGACCATTACTGAGCTGGC 751
QY 180 aAspAspIleCysCysLeu-LeuLysGluProCysValleuGlnArgAlaGlyProLeuP 200
Db 752 GGATGACATGTGCTGCTANTCTCAGGACCCCTGTGTCTGCAAAAGGCTGCCCGCTCC 811
QY 200 roGlyLysAspIle-ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGlu 219
Db 812 CTGGCAAGATATACCCCTACTCTGTGACTGTGCAAAAGACACCACTCACTGGGAAGAG 871
QY 220 LeuAspSerSerLeuLeu---PheSerGluAlaAlaThrGlyGluGluSerLeuLeuSer 238

Db 872 CTGGACAGCTCCCTCTGTTTCTTGAAGCTGCCCCAGGGGAGGAGTCTTCTCAT 931

QY 239 GUGLY--LeuArgGluSerLeuSerPheTyrIleSerLeu-AsnAspGluAla 255

Db 932 GAGGGTCTTCCGGAGTCCCTCAGCTTCTACATCAACCTGTATGACCAGGCT 985

RESULT 4

AK088672 2637 bp mRNA linear HTC 03-APR-2004

LOCUS AK088672

DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430023D24 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.

ACCESSION AK088672

VERSION AK088672.1 GI:26353729

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

MEDLINE 28499374

PUBMED 11076861

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 28499374

PUBMED 11076861

REFERENCE 6

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kaga, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Muramatsu, M. and Hayashizaki, Y.

JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of

physical and chemical research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

FEATURES

source

1. 2637

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="FANTOM DB:E430023D24"

/db_xref="taxon:10090"

/clone="E430023D24"

/cell_type="thymic cells"

/tissue_type="thymus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="2 days neonate"

358. 1137

/note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTRI)A138196, evidence: FASTV, 100%ID, 100%length, match=777)

putative"

/codon_start=1

/protein_id="BAC40495.1"

/db_xref="GI:26353730"

/translation="MGSLSSRGKTSPPSSGDPQEPVSMQPERHKVTAVALGSPDA GEQARLSRLGEPLTITISDGDWTVQSEVSGREYHMPVYAKVAHWLVEGSRREK ABEILLPGNPGAPLIRISQTRGCVSLVRLSPASWDRIRHRIQLDNGWLYIS PLTFPSLHALVEHSELDGICCPLEPCVLQKLGPLPKDTPPPVTPTSSLNWKK LDRSLFLFAPASGEASLLSEGLRESLSYSISLAEPLDDA"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

2.02e-91 2637

1032.00 209

85.88% 16

79.77% 33

76.39% 4

3 Gaps: 3

US-10-043-649-2 (1-261) x AK088672 (1-2637)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20

Db 358 ATGGGAAGTTTGTCCAGCAGAGGGAACC--TCCAGCCCCAGCCCCAGCTCTGTGT 414

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40

Db 415 CCAGACCAGGAACCCGTGTCCATGACACGAAAGACACAAAGGTCTACAGCTGTGCCCTG 474

QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60

Db 475 GGCAGTTTCCAGCAGGTGAACAGGCCAGACTATCTGTGAGACTCGGGAGCCGCTGACC 534

QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80

Db 535 ATCATCTCTGAGATGAGATTGTGTGACAGTCCAGTCCGAAGTCTCAGGACAGAGTAC 594

QY 81 AsnIleProSerValHisValAlaIysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100

.....

Pred. No.: 6.05e-91 Length: 2974
Score: 1028.00 Matches: 208
Percent Similarity: 85.88% Conservative: 17
Best Local Similarity: 79.39% Mismatches: 33
Query Match: 76.09% Indels: 4
DB: 3 Gaps: 3

US-10-043-649-2 (1-261) x AK030877 (1-2974)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 232 ATGGGAAGTTGTCCAGACAGAGGAAACC--TCCAGCCCCAGCCCCAGCTCTCTGGT 288
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 289 CCAGACCAGGAACCCGTGTCCATGCACACCAAGAACACAAAGTCAACAGCTGTGGCCCTG 348
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 349 GGCAGTTTCCAGCAGGTGAACAGGCCAGACTATCTTGAGACTCGGGAGCCGCTGACC 408
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 409 ATCATCTCTGAGGATGAGATTGTGTGACAGATCCAGTCGGAAGTCTCAGGCAGAGAGTAC 468
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 469 CACATGCCCAGTGTGTATGTGGCTAAAGTCGCCACGGGTGGCTGTACGAGGGCTGAGC 528
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 529 CGGGAGAAAGCCGAGAACTACTCTGTACTCTGGGAACCCCGAGGGCTTCTCATC 588
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 589 CGGGAGAGCCAGACCAAGAGAGGCTACTATTCCCTGTCCGCTCGACTCAGCCGCTGCA 648
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 649 TCTTGGGACCGGATCAGACACTACAGGATACAGCGCTTGTGACAAATGGCTGTACATC 708
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 709 ACACCTCGCCTCACCTTCCCTCACTCCACGCGCTTGTGGAGCATTACTCTGAGCTAGCA 768
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 769 GATGGCATCTGCTGTCCCTCAGGAGCGGTGTCTCTGACAGAGCTTGGGCCACTACCT 828
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 829 GGCAGAGATACACCTCCACCTGTGACTGTGCCAACATCATCACTAAATTGAAAAAGCTG 888
QY 221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239
Db 889 GACCGCAGCCTCTCTGTTCTTGAAAGCACTGCGAGTGGGAGGAGCATCTGCTCAGTAG 948
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 949 GGGCTCCGAGAGTCCCTCAGTTCTCTACATCAGCTTGCTGAGGAC-----CCCTTGAT 1002
QY 260 AspAla 261
Db 1003 GATGCT 1008

RESULT 6
AK020837 926 bp mRNA linear HTC 03-APR-2004
LOCUS AK020837
DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930009E21 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING
MARS, full insert sequence.
ACCESSION AK020837
VERSION AK020837.1 GI:12861542

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-2000) Yoshinide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further
details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGATCCAGAGAGCTTTTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase

and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot = 20.0 and subtraction to Rot = 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCGAGTTAATTAATTAATTCACCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.

FEATURES

source

1..926 location/Qualifiers

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A930009E21"
/clone="A930009E21"
/tissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
128..667
/note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTR)|AAL38196, evidence: FASTY, 100%ID, 100%length, match=777)
putative"

CDS

/protein_id="BAB32223.1"
/db_xref="GI:12861543"
/translation="MPSVYVAKVAHWLYEGLSREKAEILLPDPGAFILRESQT
RRCYSLSVRLSRPASWDRIHRHRIQLDNGWLITSPRLTFPSLHLYEHYSELADGI
CCPLREPCVLOKLGPLGKDTPTPTVPTSSLNWKKLDRSLFLPADASGEASLSSEG
LRESLSYISLAEDPLDDA"
907..912
/note="putative"
926
/note="putative"
/note="putative"

polyA_signal

polyA_site

ORIGIN

Alignment Scores:
Pred. No.: 1.13e-82 Length: 926
Score: 937.50 Matches: 184
Percent Similarity: 88.74% Conservative: 13
Best Local Similarity: 82.88% Mismatches: 22
Query Match: 69.39% Indels: 3
DB: 3 Gaps: 2

US-10-043-649-2 (1-261) x AK020837 (1-926)

QY 41 GlySerPheProAlaGlyGlyProAlaGluSerLeuArgLeuGlyGluProLeuThr 60
DB 5 GGCAGTTTCCAGCAGGTGACAGCGCCAGACTATCTTGAGACTCGGGGAGCCGCTGACC 64
QY 61 IleValSerGluAspGlyAspTyrTrpThrValLeuSerGluValSerGlyArgGluTyr 80
DB 65 ATCATCTCTGAGATGAGATTGTGTGACAGTCCAGTCGGAAGTCTCAGGAGAGAGTAC 124
QY 81 AsnIleProSerValHisValAlaValSerHisGlyTyrLeuTyrGlyGlyLeuSer 100
DB 125 CACATGCCCACTGTATGTGGCTAAAGTCGCCACGGGTGGCTGTACGAGGCGCTGAGC 184
QY 101 ArgGluValAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
DB 185 CGGGAGAAAGCCGAGGAAGTACTCTGTTAAGTGGGAAAGCCGAGGGGCGCTTCTCATC 244
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 245 CGGGAGAGCCAGACAGAGAGAGGCTGCTATTCCTGTCGACTCAGCCGCCCTGCA 304

QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
DB 305 TCTTGGGACCGGATCAGACACTACAGATACAGCGTCTTGACATGGCTGCTATCATC 364
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
DB 365 TCACCTCGCCTCACCTTCCCTCACTCCACGCCCTGTGTGAGCATTAATCTGAGCTAGCA 424
QY 181 AspAspIleCysCysLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB 425 GATGGCATCTGCTGTGCTCCCTCAGGAGCCGCTGTGTCTGCAAGCCTTGGCCACTACCT 484
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeu 220
DB 485 GGCATAAGATACACCTCCACCTGTGACTGTGCCAACATCATCAATAATTGAAAAAGCTG 544
QY 221 AspSerSerLeuLeuPheSerGluAla--AlaThrGlyGluGluSerLeuLeuSerGlu 239
DB 545 GACCGCAGCCTCCTGTTTCTGGAAGACACCTGCGAGTGGGAGGACATCTGTCTCAGTGA 604
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
DB 605 GGGCTCCGAGAGTCCCTCAGTCTCCTACATCAGCCCTGGCTGAGAGC-----CCCTTGAT 658
QY 260 AspAla 261
DB 659 GATGCT 664

RESULT 7
BY742155 660 bp mRNA linear EST 17-DEC-2002
LOCUS BY742155
DEFINITION BY742155 RIKEN full-length enriched, adult retina Mus musculus cDNA.
clone A930009E21 5', mRNA sequence.
ACCESSION BY742155
VERSION BY742155.1 GI:27167493
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 660)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempole, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
JOURNAL MEDLINE
PUBMED
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sueniro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES
source

Location/Qualifiers
1. 660
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A930009E21"
/tissue_type="retina"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult retina"
/note="Site 1: Salt; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5'GAGAGAGAGATCTCGAGTTAATTAAATTATCCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
PhiC1. -Retina RNA was provided by Stefano Gustincich,
Department of Neurobiology, Harvard Medical School, 220
Longwood Ave., Boston, MA02115, USA, whose assistance we
gratefully acknowledge."

ORIGIN
Alignment Scores:

Pred. No.:
Score:
Percent Similarity:

2.07e-77
882.50
85.97%

Length: 660
Matches: 175
Conservative: 15

Best Local Similarity: 79.19%
Query Match: 65.32%
DB: 6

Mismatches: 24
Indels: 8
Gaps: 2

us-10-043-649-2 (1-261) x BY742155 (1-660)

QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 5 GGCAGTTTCCCGACGAGGTGAACAGGCCAGACTATCTTGAGACTCGGGAGCCGCTGACC 64
QY 61 ILeValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 65 ATCATCTTGAGGATGAGATTGGTGACACGTCCAGTCGGAAGTCTCAGGACAGAGTAC 124
QY 81 AsnIleProSerValHisValAlaIleValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 125 CACATGCCAGTGTGTATGTGGCTAAAGTCGCCACGGGTGGTGTACGAGGGCTGAGC 184
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 185 CGGAGAAAGCCGAGAACTACTCTGTATTACCTGGGAACCCGGAGGGCTTCTCATC 244
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 245 CGGAGAGCCAGACAGAGAGGCTGTATTTCCTGTCCGACTCAGCCGCTGCA 304
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 305 TCTTGGACCGGATCAGACACTACAGATACAGCGCTTGACAAATGGCTGGCTGTACATC 364
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 365 TCACCTCGCTCACCTTCCCTTCACCTCCACGCGCTGGTGAGCACTTACTGTAGCTAGCA 424
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 425 GATGCATCTGCTGTGCTCCCTCAGGAGACCGGTGTCTGTCTGCAAGACTTGGCCACTACT 484
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 485 GGCAGATACACCTCCACCTGTGACTGTGCCACATCATCAATAATGGAAGAAAGCTG 544
QY 221 AspSerSerLeuLeuPheSerGluAla--AlaThrGlyGluGluSerLeuSerGlu 239
Db 545 GACCGTAGCTCTCTGTTCTTGGAAGCACTGCCAAGTGGGAGGAGCATCTGCTCATTTAG 604
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 605 GGGCTCTCAGAGTC-CTCACCTTTTACTC-----ATCCCGTTGAT 645
QY 260 Asp 260
Db 646 GAC 648

RESULT 8
AL541041

LOCUS 1032 bp mRNA linear EST 24-MAR-2004
DEFINITION AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23

5-PRIME, mRNA sequence.

ACCESSION AL541041

VERSION AL541041.3 GI:45716635

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:30544829.
Contact: Genoscope
Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 9825.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DE005AF12QPI&c=9825.r.
Location/Qualifiers
1. 1032

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YK23"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Alignment Scores:
Pred. No.: 5.59e-76 Length: 1032
Score: 871.00 Matches: 169
Percent Similarity: 96.02% Conservative: 0
Best Local Similarity: 96.02% Mismatches: 7
Query Match: 64.47% Indels: 0
DB: 1 Gaps: 0

US-10-043-649-2 (1-261) x AL541041 (1-1032)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
Db 395 ATGGGAAGTCTGCCCGACAGAAATCTCTGCCAGCCCAAGCTTGAGTCTCTGTC 454
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLyAlaThrAlaValAlaLeu 40
Db 455 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAAAGCCACAGCCGTGGCCCTG 514
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 515 GGCAGTTCCTCCGCGACGTGGCCCGCGAGCTGTGCTGAGACTCGGGGACCATTTGACC 574
QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 575 ATCGTCTGTGAGATGAGACTGTGTGACGGTGTCTGAAGTCTCAGGCAGAGATAT 634
QY 81 AsnIleProSerValHisValAlaValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 635 AACATCCCCCAGCGTCCACGTGCCAAAGTCTCCCATGGGTGCTGTATGAGGCCCTGAGC 694
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 695 AGGAGAAAGCAGAGAACTGCTGTGTACTCTGCGAACCCCTGGAGGGGCTTCTCATC 754
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 755 CGGAGAGCCAGAMCAGAGAGAGGCTTACTCTCTGTCACTCCGCTCAAGCCCTGCA 814
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 815 TCCTGGGACCGCATCAGACACTAMAGATCCACTGCTTGACAAATGGCTGGCTGATCATC 874
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyr 176
Db 875 TCACGGCGCCTTAACTTCCCTCACTCCAGGCCCTGTGTGACMATTAAC 922

RESULT 9
BG284179

LOCUS BG284179 566 bp mRNA linear EST 21-FEB-2001
DEFINITION 602408226F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520382 5',
mRNA sequence.
ACCESSION BG284179
VERSION BG284179.1 GI:13034866
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 566)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10418 row: C column: 07
High quality sequence start: 2
High quality sequence stop: 566.
Location/Qualifiers
1. 566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4520382"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN
Alignment Scores:
Pred. No.: 2.02e-60 Length: 566
Score: 711.00 Matches: 136
Percent Similarity: 99.27% Conservative: 0
Best Local Similarity: 99.27% Mismatches: 1
Query Match: 52.63% Indels: 0
DB: 4 Gaps: 0

US-10-043-649-2 (1-261) x BG284179 (1-566)

QY 125 ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144
Db 4 ACGCGTCCGGGCTTACTCTCTGTGATCCGCTCAGCCGCTGCAATCTGGGACCGG 63
QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeu 164
Db 64 ATCAGACACTACAGATCACTGCTTGAACAATGGCTGTGACTGATCACTCAAGCCGCTC 123
QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspIleCys 184
Db 124 ACCCTCCCTCACTCCAGGCCCTGTGTGACCATTAATCTGAGCTGGCGATGACATCTGC 183
QY 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204
Db 184 TGCTTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCCGCTCCCTGCAAGATATA 243
QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224
Db 244 CCCCTACTGTGACTGTGACAGAGACCACTCAACTGGAAGAGCTGGAAGCTCCCTC 303
QY 225 LeuPheSerGluAlaIleThrGlyGluGluSerLeuLeuSerGluGlyLeuArgGluSer 244

Db	304	CTGTTTTCGAAGCTGCCACAGGGGAGAGTCTTCTCAGTGAGGGTCTCCGGAGTCC	363
Qy	245	leuSerPheTyrtlleSerleuAsnaspGluaIaValSerleuAspaspAla	261
Db	364	CTCAGCTTCTACATCAAGCCTGAATGACGAGGCTGTCTTCTTGATGATGCC	414
RESULT 10			
BQ054265			
LOCUS			
DEFINITION	BQ054265	986 bp	mRNA linear EST 29-MAR-2002
ACCESSION	AGENCOURT_6830248	NIH_MGC_106	Homo sapiens cDNA clone IMAGE:5936339
VERSION	BQ054265		
KEYWORDS	BQ054265.1	GI:19813605	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1	(bases 1 to 986)	
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-r@mail.nih.gov		
	Tissue Procurement: Dr. Daniel McVicar, DBS/NCI		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: LLCM2125 row: i column: 12		
	High quality sequence stop: 515.		
FEATURES	Location/Qualifiers		

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936339"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/notes="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

```

ORIGIN

Alignment Scores:

Pred. No.:	2e-57
Score:	684.50
Percent Similarity:	73.57%
Best Local Similarity:	72.25%
Query Match:	50.67%
DB:	5

Length:	986
Matches:	164
Conservative:	3
Mismatches:	16
Indels:	44
Gaps:	3

US-10-043-649-2 (1-261) X BQ054265 (1-986)

QY	1	MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal	20
Db	279	ATGGGAAGTCTGCCACGACAGAGAAATCTCTGCCAAGCCCAAGCTTAGTTCTCTGTC	338
QY	21	GIINGlyGIINGlyProValThrMetGluAlaGluArgSerLySAlaThrAlaValAlaLeu	40
Db	339	CAAGGCCAGGGAACCTGTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTCCTG	398
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60

Db	399	GGCAGTTTCCCGCAGGTGGCCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC	458
QY	61	IleValSerGluAspGlyAspTyrTrpThrValLeuSerGluValSerGlyArgGluTyr	80
Db	459	ATCGTCTGTGAGATGGAGACTGGTGACGGTGCTGTCTGAAGTCTCAGCCAGAGATAT	518
QY	81	AsnIleProSerValHisValAlaIysValSerHisGlyTyrLeuTyrGluGlyLeuSer	100
Db	519	AACATCCCCACGCTCCACGCTGGCCAAAGTCTCCATGGTGCTGTATGAGGGCCTGAGC	578
QY	101	ArgGluIysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAla-PheLeuIle	120
Db	579	AGGAGAAAGCAGAGAACTGCTGTGTAACTGGGAACCTGAGGGGCCCTTCCTCAT	638
QY	120	eaArgGlu-SerGlnThrArgArgGlySerTyrSerLeuSerVal-ArgLeuSerArg-Pr	139
Db	639	CCGGGAGAGCCAGACCAGAGAGAGGCTCTTACTCTGTGTCAGTCCCGCCTCAGCCGCC	698
QY	139	oAlaSerTyr-AspArg-IleArgHisTyrArg-IleHisCysLeuAspAsnGlyTyrIle	158
Db	699	TGCATCCTGGGGACCGGGATCAGACCCTACAGGGATTCGCCCTTGAAACCATTTGGCT	758
QY	158	u-----TyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAs	174
Db	759	TGGCTGTTAATTTTAAACCGGGGCTTACCTTTTCCC-----	798
QY	174	pHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeuGlu	194
Db	799	-----CTTAA	803
QY	194	naArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPr	214
Db	804	ATTTCAGGGGCGCCCTTGGGGGGAACCATTTT-----ACTCC	839
QY	214	oLeuAsn	216
Db	840	TTTAAAC	846

Dib

RESULT 11	
BU944126	
LOCUS	960 bp mRNA linear EST 18-OCT-2002
DEFINITION	AGENCOURT 10545003 NIH MGC 107 Homo sapiens CDNA clone
IMAGE:6728350 5',	mRNA_sequence.

DEFINITION

ACCESSION	BU944126
VERSION	BU944126.1
KEYWORDS	GI.24132945
SOURCE	EST.
	Homo sapiens (human)

ORGANISM

Homo sapiens (human)

ORGANISM

ORGANISM Homo sapiens
Eumaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.

REFERENCE

REFERENCE
1 (bases 1 to 580),
NTH-MGC b110://mac.nci.nih.gov/.

ADULTS
ATTENDING

Unpublished (1999)

COMMENT

Email: cgapbs-r@mail.nih.gov

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The J.M.A.S.: Consortium (2002)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

<http://image.11n1.gov>

High quality sequence stop: 628.

РЕЗУЛЬТАТЫ

FEATURES	Location/Qualifiers
1	860

Sou

/organism="Homo sapiens"

```
/db xref="taxon:9606"
```

```
/tissue type="adenocarcinoma, cell line"
```

```
/lab_host="DHLUB (phage-resistant)"
```

/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 3.61e-50 Length: 960
Score: 611.50 Matches: 124
Percent Similarity: 86.71% Conservative: 0
Best Local Similarity: 86.71% Mismatches: 1
Query Match: 45.26% Indels: 18
DB: 5 Gaps: 1

US-10-043-649-2 (1-261) x BU944126 (1-960)

QY 120 IleaArgLuserGlnThrArgArgLysSerTyrSerLeuSerValArgLeuSerArgPro 139
Db 3 ATCCGGAGAGACGACAGACGAGAGGCTTACTCTCTGTGAGTCCGCCCTCAGCCGCTTAC 62
QY 140 AlaSerTrpAspArgLysArgHisTyrArgLysHisCysLeuAspAsnGlyTrpLeuTyr 159
Db 63 GCATCTGGGACCGGATCAGACTACAGATCCACTGCTTGACAAATGGCTGGCTGTAC 122
QY 160 IleserProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeu 179
Db 123 ATCTCACCGCGCCTCACCCTTCCCTCACTCCAGGCCCTGTGGAGCCATTAC----- 173
QY 180 AlaAspAspIleCysCysLeuLeuLysGluProCysValleuGln-ArgAlaGlyProle 199
Db 174 -----TCTGAGGGCTGGCCGCT 191
QY 199 uProGlyLysAspIleProleuProValThrValGlnArgThrProleuAsnTrpLysG 219
Db 192 CCTGGCAAGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGGAAAGA 251
QY 219 uLeuAspSerSerLeuLeuPheSerGlnAlaAlaThrGlyGluGluSerLeuSerG 239
Db 252 GCTGACAGACTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTTCTCAGTGA 311
QY 239 uGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeu 259
Db 312 GGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTTTGGA 371
QY 259 pAspAla 261
Db 372 TGATGCC 378

RESULT 12
BQ053486 878 bp mRNA linear EST 29-MAR-2002
LOCUS BQ053486
DEFINITION AGENCOURT_6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253
5', mRNA sequence.
ACCESSION BQ053486
VERSION BQ053486.1 GI:19812826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 878)
NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM2122 row: 1 column: 06
High quality sequence stop: 394.
Location/Qualifiers
1..878

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5935253"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 1.52e-45 Length: 878
Score: 564.50 Matches: 135
Percent Similarity: 64.04% Conservative: 11
Best Local Similarity: 59.21% Mismatches: 20
Query Match: 41.78% Indels: 64
DB: 5 Gaps: 5

US-10-043-649-2 (1-261) x BQ053486 (1-878)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 273 ATGGAGTCTGCCAGACAGAAAGAAATCTTGGCAAGCCCAAGCTTGATTCCTGTGC 332
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 333 CAAGGCCAGGACTGTGACCATGGAAGCAGAGAGAAAGCAAGGCCACAGCCGTGCCCTG 392
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProleuThr 60
Db 393 CGCAGTTTCCCGCAGAGTGGCCCGGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 452
QY 61 ILeValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 453 ATCGTCTGAGAGTGAAGACTGTGAGCGGTGCTGTGAATCTCAGCAGAGAGTAT 512
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 513 AACATCCCATCGTCCACGTGGCCAAAGTCTCCATGCGTGGCTGATGAGGGCT--GAN 570
QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 571 CAAGAAACCAAGGAGACTGTGTGTAACTGGGAAACCTGAGGGGCTTCTCTCCTC 630
QY 121 -ArgGluSerGlnThrArgArgLysSer---TyrSerLeuSerValArgLeuSerArgPr 139
Db 631 CCGGAGAACAGACAGAGAGAGGCTCTTAACCTCTGTGTA-GTCCGCTTAGCCGCC 689
QY 139 oAlaSerTrpAspArgLysArgHisTyrArgLysHisCysLeuAspAsnGlyTrpLeuTyr 159
Db 690 C-----
QY 159 rIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLe 179
Db 690 -----
QY 179 uAlaAspAspIleCysCysLeuLeuLysGluProCysValleuGlnArgAlaGlyProle 199

Db 691 -----TGATTCCTGGGGAACCG-----GATTGGAACCTTA 722
QY 199 uProGlyLysAspIleProLeuProValThr-----Va 210
Db 723 CCCGGGACCC-----CCCTGGCCTTGACAAATGGGGCTTGGGCTTTGGAAATTTTTT 776
QY 210 lGlnArgThrProLeuAsnTrp 217
Db 777 ACACCGGGGGCCCTTAAACTTT 798

RESULT 13
AA959151 377 bp mRNA linear EST 08-MAY-1998
AA959151
LOCUS v251906.r1 Soares thymus_2NbMT Mus musculus cDNA clone
DEFINITION IMAGE:1330042 5' similar to TR:Q13239 Q13239 PUTATIVE SRC-LIKE
ADAPTER PROTEIN ; mRNA sequence.
ACCESSION AA959151 GI:3124344
VERSION AA959151.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 377)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Stepcoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The Washu-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:689586
Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
source
location/Qualifiers
1..377
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1330042"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares thymus_2NbMT"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer (5'
TGTTACCAATCTGAAGTGGAGCGCGCTTTTCTTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

ORIGIN
Alignment Scores:
Pred. No.: 9.8e-46 Length: 377
Score: 561.00 Matches: 110
Percent Similarity: 93.50% Conservative: 5
Best Local Similarity: 89.43% Mismatches: 8
Query Match: 41.52% Indels: 1

DB: 1 Gaps: 0
US-10-043-649-2 (1-261) x AA959151 (1-377)

QY 40 lGluGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyProLeu 59
Db 8 CTGGGAGATTTCACGACGAGTGACAGAGCCAGACTATCTTGAGACTCGGGAGCCGCTG 67

QY 60 ThrIleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGlu 79
Db 68 ACCATCATCTCTGAGATGAGATTGGTGGACAGTCCAGTCCGGAAGTCTCAGGACAGAG 127

QY 80 TyrAsnIleProSerValHisValAlaValSerHisGlyTrpLeuTyrglyGlyLeu 99
Db 128 TACCACATGCCCGAGTGTATGTGCTAAAGTCGCCACCGGTGGCTGTACGAGGCTG 187

QY 100 SerArgGlyLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeu 119
Db 188 AGCCGGAGAAAGCCGAGAACTACTCCTGTACTGGGAAACCCCGAGGTCA-TTCCTC 246

QY 120 lIleArgGlySerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgPro 139
Db 247 ATCCGGAGAGCCAGACCAAGAGAGGCTGCTATTCCTGTCCTCCGACTCAGCCGCT 306

QY 140 AlaSerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyr 159
Db 307 GCATCTTGGACCGGATCAGACACTACAGGATACAGCGCTTGACATATGCTGCTGTAC 366

QY 160 lIleSerPro 162
Db 367 ATCTCACCCT 375

RESULT 14
BB635615 660 bp mRNA linear EST 26-OCT-2001
LOCUS BB635615 RIKEN full-length enriched, 0 day neonate thymus Mus
DEFINITION musculus cDNA clone A430076N07 5', mRNA sequence.
ACCESSION BB635615 GI:16471660
VERSION BB635615.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 660)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Komno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
TITLE Yoshihide Hayashizaki
JOURNAL Contact: Yoshihide Hayashizaki
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

Source Location/Qualifiers

1..660

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A430076N07"

/tissue_type="thymus"

/dev_stage="0 day neonate"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 0 day neonate
thymus"

/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCCAAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

ORIGIN

Alignment Scores:

Pred. No.: 2.51e-45 Length: 660
Score: 560.50 Matches: 115
Percent Similarity: 87.23% Conservative: 8
Best Local Similarity: 81.56% Mismatches: 17
Query Match: 41.49% Indels: 1
DB: 2 Gaps: 1

US-10-043-649-2 (1-261) x BB635615 (1-660)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 241 ATGGGAAGTTGTCCAGCAGAGAGGAAACC--TCCAGCCCCAGCCCCAGCTCTCTGT 297
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 298 CCAGACCAGGAACCGGTGTCCATGCAACCAAGACACACAAGGTCAAGCTGTGGCCCTG 357
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 358 GGCAGTTTCCAGCAGGTGAACAGGCCAGACTATCTTGAGACTCGGGGAGCCGCTGACC 417
QY 61 IleValSerGluAspGlyAspTyrTyrThrValLeuSerGluValSerGlyArgGluTyr 80
Db 418 ATCATCTCTGAGAGATGAGATTGTGTGACACTCCAGTCGGAAGTCTCAGGACAGAGATAC 477
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 478 CACATGCCCAAGTGTATGTGGCTAAAGTCCGCCACGGGTGGTGTACGAGGGCCTGAGC 537

QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 538 CCGGAGAAAGCCGAGAGACTACTCTGTACTCTGNGAACCCCGAGGGCCCTTCTCATC 597
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 598 CCGGAGAGCCAGACACAGAGAGAGGCTGCTATTCCCTGTCGTTGACTCAGCCGCCCTGCA 657
QY 141 Ser 141
Db 658 TCT 660

RESULT 15

AL844311

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AL844311 597 bp mRNA linear EST 30-JUL-2002
AL844311 pool_YT_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
AL844311.1 GI:22019093
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens EST sequence
Unpublished (2002)
Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name : scdd10818.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool_YT_11b v SPD cDNA library. Further information can be found at
<http://www.sanger.ac.uk/Teams/Team69/>.

FEATURES

Source Location/Qualifiers

1..597

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/map="20"

/clone_lib="pool_YT_11b_v_SPD"

ORIGIN

Alignment Scores:

Pred. No.: 7.64e-45 Length: 597
Score: 555.00 Matches: 107
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.08% Indels: 0
DB: 1 Gaps: 0

US-10-043-649-2 (1-261) x AL844311 (1-597)

QY 155 AsnGlyTyrPheLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp 174
Db 13 AATGCTGGCTGTACATCTCACACGCTCACTTCCCTCCTACCTCAGGCGCTGTGAC 72
QY 175 HisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeuGln 194
Db 73 CATTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAAGAGCCCTGTGCTGAC 132
QY 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro 214
Db 133 AGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACCTGTGACTGTGACAGACACCA 192

QY	215	LeuAsnTrpIysGluLeuAspSerSerLeuPheSerGluAlaIatHrgIYGlU	234
Db	193	CTCAACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAG	252
QY	235	SerLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu	254
Db	253	TCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAG	312
QY	255	AlaValSerLeuAspAspAla	261
Db	313	GCTGTCTCTTTGGATGATGCC	333

Search completed: November 17, 2004, 03:11:31
Job time : 3383 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2004, 20:44:19 ; Search time 3247 Seconds
(without alignments)
8820.951 Million cell updates/sec

Title: US-10-043-649-1
Perfect score: 786

Sequence: 1 atcggaagtcgtccacgacag.....tctcttgatgatgcctag 786

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 segs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	734.6	93.5	1002	5	BQ052308	BQ052308 AGENCOURT
2	654	83.2	1020	5	BQ054281	BQ054281 AGENCOURT
3	645	82.1	1069	5	BQ052468	BQ052468 AGENCOURT
4	537.8	68.4	2637	3	AK088672	AK088672 Mus muscu
5	536.2	68.2	2974	3	AK030877	AK030877 Mus muscu
6	529	67.3	1032	1	AL541041	AL541041 AL541041
7	478	60.8	926	3	AK020837	AK020837 Mus muscu
8	453.8	57.7	660	6	BY742155	BY742155 BY742155
9	405	51.5	566	4	BG284179	BG284179 602408226
10	386.2	49.1	986	5	BQ054265	BQ054265 AGENCOURT
11	331.2	42.1	878	5	BQ053486	BQ053486 AGENCOURT
12	326.8	41.6	778	4	BG178487	BG178487 602328305
13	322.4	41.0	597	1	AL844311	AL844311 AL844311
14	321	40.8	960	5	BU944126	BU944126 AGENCOURT
15	317.4	40.4	614	1	AL844307	AL844307 AL844307
16	314.6	40.0	606	1	AL844309	AL844309 AL844309
17	296.2	37.7	660	2	BB635615	BB635615 BB635615
18	289.8	36.9	762	5	BP162888	BP162888 BP162888
19	283.6	36.1	377	1	AA959151	AA959151 vz51g06.r
20	283.2	36.0	701	7	CK833360	CK833360 4057315 B
21	283	36.0	781	7	CK596391	CK596391 AGENCOURT
22	277.6	35.3	569	6	CB426333	CB426333 601508 MA
23	275.6	35.1	627	2	BB619854	BB619854 BB619854
24	261.4	33.3	794	4	BG677567	BG677567 602624118

25	249.4	31.7	656	7	CN792083	CN792083 4126904 B
26	221.6	28.2	603	5	BQ553005	BQ553005 H4019E02-
27	220	28.0	670	1	AI510095	AI510095 m143c04.Y
28	215.8	27.5	640	5	BQ553006	BQ553006 H4019E02-
29	196.2	25.0	575	5	BP156169	BP156169 BP156169
30	175.6	22.3	701	5	BX849096	BX849096 BX849096
31	170.8	21.7	569	6	CB514688	CB514688 ssalrpb52
32	166.8	21.2	642	1	AL844308	AL844308 AL844308
33	163.4	20.8	831	9	AY416277	AY416277 Homo sapi
34	162.2	20.6	871	5	BQ436143	BQ436143 AGENCOURT
35	161.8	20.6	619	1	AL844312	AL844312 AL844312
36	160.6	20.4	570	1	AI471720	AI471720 tal6a01.x
37	160.6	20.4	1762	3	CR617843	CR617843 full-leng
38	160.2	20.4	882	5	BX415149	BX415149 BX415149
39	160.2	20.4	1015	5	BX436423	BX436423 BX436423
40	160.2	20.4	1037	1	AL539427	AL539427 AL539427
41	160.2	20.4	1054	1	AL549826	AL549826 AL549826
42	159.8	20.3	966	1	AL551370	AL551370 AL551370
43	159.2	20.3	1063	5	BQ072745	BQ072745 AGENCOURT
44	159	20.2	972	5	BQ707614	BQ707614 AGENCOURT
45	158	20.1	562	5	BX383606	BX383606 BX383606

ALIGNMENTS

RESULT 1
BQ052308
LOCUS
DEFINITION AGENCOURT_6868571 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933542
5', mRNA sequence.
ACCESSION BQ052308
VERSION BQ052308.1 GI:19811648
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1002)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2118 row: d column: 23
High quality sequence stop: 670.

FEATURES

source

1. 1002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933542"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match

93.5%; Score 734.6; DB 5; Length 1002;

Best Local Similarity 98.3%; Pred. No. 4.8e-183; Matches 774; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

QY	1	ATGGGAAGTCTGCCAGCAGAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC	60
Db	96	ATGGGAAGTCTGCCAGCAGAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC	155
QY	61	CAAGGCCAGGACTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG	120
Db	156	CAAGGCCAGGACTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG	215
QY	121	GGCAGTTTCCCGGAGGTGGCCCGGCGAGCTGTCCCTGAGACTCGGGGAGCCATTGACC	180
Db	216	GGCAGTTTCCCGGAGGTGGCCCGGCGAGCTGTCCCTGAGACTCGGGGAGCCATTGACC	275
QY	181	ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAAGAGTAT	240
Db	276	ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAAGAGTAT	335
QY	241	AACATCCCCCAGCGTCCACGTGGCCCAAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC	300
Db	336	AACATCCCCCAGCGTCCACGTGGCCCAAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC	395
QY	301	AGGAGAAAGCAGAGAACTGCTGTGTACCTGGAAACCTTGAGGGGCTTCTCTCATC	360
Db	396	AGGAGAAAGCAGAGAACTGCTGTGTACCTGGAAACCTTGAGGGGCTTCTCTCATC	455
QY	361	CGGAGAGCCAGACCAAGAGAGGCTTCTACTCTGTGACGTCGCTCAGCCGCTGCA	420
Db	456	CGGAGAGCCAGACCAAGAGAGGCTTCTACTCTGTGACGTCGCTCAGCCGCTGCA	515
QY	421	TCCTGGGACCCGATCAGACTACAGATCCACTGCCTGACATGGCTGCTGATC	480
Db	516	TCCTGGGACCCGATCAGACTACAGATCTACTGCCTGACATGGCTGCTGATC	575
QY	481	TCACCGCGCTCACTTCCCTCACTCCAGGCGCTGTGACCATTAATCTGAGCTGGCG	540
Db	576	TCACCGCGCTCACTTCCCTCACTCCAGGCGCTGTGACCATTAATCTGAGCTGGCG	635
QY	541	GATGACATCTGCTGCTCACTCAAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCT	600
Db	636	GATGACATCTGCTGCTCACTCAAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCCT	695
QY	601	GGCAAGATATACCCCTACTGCTGAGTGTGACAGAGACCACTCACTGAAAGAGCTG	660
Db	696	GGCAAGATATACCCCTACTGCTGAGTGTGACAGAGACCACTCACTGAAAGAGCTG	755
QY	661	GACAGTCTCTCTCTGTTTCTGAAGCTGCCACA-GGGAGAGAGTCTTCTCAGTGAGG	719
Db	756	GACAGTCTCTCTCTGTTTCTGAAGCTGCCACAAGGGAGAGAGTCTTCTCAGTGAGG	815
QY	720	--TCTCCGGAGATCCCTCAGCTTCTACATCAGCCTGAATGACGA-GGCTGTCTCTTGA	776
Db	816	GTTCTCCGGAGATCCCTCAGCTTCTACATCAGCCTGAATGACGAAGGCTGTCTTGA	875
QY	777	TGATGCC 783	
Db	876	TGATGCC 882	

RESULT 2
BQ054281 1020 bp mRNA linear EST 29-MAR-2002
LOCUS BQ054281
DEFINITION AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
5', mRNA sequence.
ACCESSION BQ054281
VERSION BQ054281.1 GI:19813621
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1020)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2125 row: j column: 11
High quality sequence stop: 556.
Location/Qualifiers

FEATURES
source 1..1020
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 83.2%; Score 654; DB 5; Length 1020;
Best Local Similarity 95.4%; Pred. No. 9.9e-162;
Matches 748; Conservative 0; Mismatches 27; Indels 9; Gaps 7;

QY	1	ATGGGAAGTCTGCCAGCAGAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC	60
Db	212	ATGGGAAGTCTGCCAGCAGAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC	271
QY	61	CAAGGCCAGGACTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG	120
Db	272	CAAGGCCAGGACTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGGCCCTG	331
QY	121	GGCAGTTTCCCGGAGGTGGCCCGGCGAGCTGTCCCTGAGACTCGGGAGCCATTGACC	180
Db	332	GGCAGTTTCCCGGAGGTGGCCCGGCGAGCTGTCCCTGAGACTCGGGAGCCATTGACC	391
QY	181	ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAAGAGTAT	240
Db	392	ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAAGAGTAT	451
QY	241	AACATCCCCCAGCGTCCACGTGGCCCAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC	300
Db	452	AACATCCCCCAGCGTCCACGTGGCCCAAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC	511
QY	301	AGGAGAAAGCAGAGAACTGCTGTGTTAATCTGGAAACCTTGAGGGGCTTCTCATC	360
Db	512	AGGAGAAAGCAGAGAACTGCTGTGTTAATCTGGAAACCTTGAGGGGCTTCTCATC	571
QY	361	CGGAGAGCCAGACCAAGAGAGGCTTACTCTGTGACGTCGCTCAGCCCTGAGCCCTGCA	420
Db	572	CGGAGAGCCAGACCAAGAGAGGCTTACTCTGTGACGTCGCTCAGCCCTGAGCCCTGCA	631
QY	421	TCCTGGGACCCGATCAGACTACAGATCCACTGCCTTGACAAATGGCTGCTGATCATC	480
Db	632	TCCTGGGACCCGATCAGACTACAGATCCACTGCCTTGACAAATGGCTGCTGATCATC	691
QY	481	TCACCGCGCTCACTTCCCTCACTCCAGG-CCCTGGTGAACCATTAATCTGAGCTGGC	539
Db	692	TCACCGCGCTCACTTCCCTCACTCCAGGCGCCCTGGTGAACCATTAATCTGAGCTGGC	751

QY 540 GGATGACATCTGCTGCCTA-CTCAGAGACCCCTGTGTCTCTGACAGGGCTGGCCGCTCC 598
Db 752 GGATGACATCTGCTGCCTA-CTCAGAGACCCCTGTGTCTCTGACAGGGCTGGCCGCTCC 811
QY 599 CTGGCAAGATATA-CCCCCTACCTGTGACTGTGCAGAGACACCACTCAACTGGAAGAG 657
Db 812 CTGGCAAGATATA-CCCCCTACCTGTGACTGTGCAGAGACACCACTCAACTGGAAGAG 871
QY 658 CTGGACAGCTCCCTCTG-TTTTCTGAAAGCTGCCACA-GGGAGAGAGTCTTCTCTGAGT 714
Db 872 CTGGACAGCTCCCTCTG-TTTTCTGAAAGCTGCCACA-GGGAGAGAGTCTTCTCTGAGT 931
QY 715 GAGGG-TCTCCGGAGTCCCTCAGCTTCTACATCAGCCTG-AATGACGAGGCTGTCTCT 771
Db 932 GAGGGTCTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGTAATGACCAAGGCTGGCTTC 991
QY 772 TTGG 775
Db 992 TTGG 995

RESULT 3
BQ052468
LOCUS 1069 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT_6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
5', mRNA sequence.
ACCESSION BQ052468
VERSION BQ052468.1 GI:19811808
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1069)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2118 row: n column: 13
High quality sequence stop: 681.
Location/Qualifiers
1. 1069

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933772"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 82.1%; Score 645; DB 5; Length 1069;
Best Local Similarity 93.4%; Pred. No. 2.4e-159;
Matches 737; Conservative 0; Mismatches 0; Indels 52; Gaps 4;
QY 1 ATGGGAAGTCTGCCAGCAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTCTGTC 60
|||||

Db 89 ATGGGAAGTCTGCCAGCAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTCTGTC 148
QY 61 CAAGGCGAGGGACCTGTGACCATGGAACAGAGAGAGAAAGCAAGCCAGCCGTGCGCTG 120
Db 149 CAAGGCGAGGGACCTGTGACCATGGAACAGAGAGAGAAAGCAAGCCAGCCGTGCGCTG 208
QY 121 GGCAGTTTCCCGCAGAGTGGCCCGCCAGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 209 GGCAGTTTCCCGCAGAGTGGCCCGCCAGAGCTGTGCTGAGACTCGGGAGCCATTGACC 268
QY 181 ATGCTCTCTGAGATGAGACTGTGAGACGGTGTCTGTAAGTCTCAGGACAGAGATAT 240
Db 269 ATGCTCTCTGAGATGAGACTGTGAGACGGTGTCTGTAAGTCTCAGGACAGAGATAT 328
QY 241 AACATCCCCAGCGCTCCACGTGGCCCAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300
Db 329 AACATCCCCAGCGCTCCACGTGGCCCAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 388
QY 301 AGGGAAGAAAGCAGAGAACTGCTGTGTTACCTGGGAACCTTGAGGGGCGCTTCCTCATC 360
Db 389 AGGGAAGAAAGCAGAGAACTGCTGTGTTACCTGGGAACCTTGAGGGGCGCTTCCTCATC 448
QY 361 CGGAGAGCCAGACAGAGAGAGCTCTTACTCTCTGTCTGATCCGCTCAGCCGCTGCA 420
Db 449 CGGAGAGCCAGACAGAGAGAGCTCTTACTCTCTGTCTGATCCGCTCAGCCGCTGCA 508
QY 421 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACATGAGTGGCTGTATCATC 480
Db 509 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGACATGAGTGGCTGTATCATC 568
QY 481 TCACCGCGCTCCTACCTTCCCTCTCACTCCAGGCCCT-GGTGACCATTAATTCTGAGTGGC 539
Db 569 TCACCGCGCTCCTACCTTCCCTCTCACTCCAGGCCCTGGGTGACCATTAATTCTGAG----- 623
QY 540 GGATGACATCTGCTCCTACTCAAGAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTGCC 599
Db 624 -----GGGCTGGCCGCTGCC 639
QY 600 TGGCAAGGATATACCCCTTACTGTGACTGTGACAGAGACACCACTCAACTGGAAGAGCT 659
Db 640 TGGCAAGGATATACCCCTTACTGTGACTGTGACAGAGACACCACTCAACTGGAAGAGCT 699
QY 660 GGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACA-GGGAGAGAGTCTCTCTCAGTGAGG 718
Db 700 GGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAAGGGAGAGTCTCTCTCAGTGAGG 759
QY 719 GTCTCCGGG-AGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGAT 777
Db 760 GTCTCCGGGAAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGAT 819
QY 778 GATGCTAG 786
Db 820 GATGCTAG 828

RESULT 4
LOCUS AK088672 2637 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
full-length enriched library, clone:E430023D24 product:MODULATOR OF
ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.
ACCESSION AK088672
VERSION AK088672.1 GI:26353729
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253

Db	1075	GGGCTCCGAGAGTCCCTCAGTTCTTACATCAGCCTGGCTGAGACCCTTGATGATGCT	1134
QY	778	GATGCCTAG	786
Db	1135	TAGCCCTGG	1143
RESULT 5			
AKO30877			
LOCUS			
DEFINITION	AKO30877	2974 bp	mRNA linear HTC 03-APR-2004
	Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830437K10 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.		
ACCESSION	AKO30877		
VERSION	AKO30877.1	GI:26326848	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE			
AUTHORS	1 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE			
AUTHORS	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Katsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE			
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE			
AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE			
AUTHORS	6 (bases 1 to 2974)		
TITLE	Adachi,U., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoaka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,U., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.		
JOURNAL	Direct Submission		

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers

FEATURES
source

1..2974
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:5830437K10"
/db_xref="taxon:10090"
/clone="5830437K10"
/sex="male"
/tissue_type="thymus"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
232..1011
/note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTR|AAL38196, evidence: FASTY, 100%ID, 100%length, match=777)
putative"
/codon_start=1
/protein_id="BAC27168.1"
/db_xref="GI:26326849"
/translation="MGSLSSRGTSSPSRSSGPDQEPVSMOPERHKVTAVALGFPA GEQARLSRLGEP LTTISEDGMWTVQSEVSGREYHMPYVAKAHGWLVEGLSREK AEEILLPGNPGGAFLIRESQTRRGYSLSVRLSRPASMDRIHRHRIQLDNGWLYIT PLRTFPSLHALVEHYSLELDGICCPLEPCVLQKLGPKDTPPVTVPPTSLSLWKK LDRSLFLLEAPASGEASLSLSEGLRESLSYSISLAEDPLDDA"

ORIGIN

Query Match	Best Local Similarity	68.2%;	Score 536.2;	DB 3;	Length 2974;	
Matches	645;	Conservative	0;	Mismatches 138;	Indels 6;	Gaps 2;
QY 1	ATGGAAGTGTGCCCCAGCAGAGAATAATCTCTGCCAAGCCCCAAGCTTGAGTTCCTCTGTC	60				
Db 232	ATGGGAAGTTTGTCCAGCAGAGGGAAC---CTCCAGCCCCAGCTCCTCTGGT	288				
QY 61	CAAGGCCAGGACCTGTGACCATGGAACAGAGAAAGCAAGGCCACAGCCGTGGCCCTG	120				
Db 289	CCAGACCCAGGAACCCGTGTCCATGCAACAGAAAGACAAAGGTCAAGCTGTGGCCCTG	348				
QY 121	GGCAGTTTCCGGCAGGTGCGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC	180				
Db 349	GGCAGTTTCCAGCAGGTGAACAGCCAGACTATCTTGAGACTCGGGAGCCGTGACC	408				
QY 181	ATCGTCTGTGAGATGAGACTGTGTGACGCGTGTCTGTAAGTCTCAGGCAGAGATAT	240				
Db 409	ATCATCTCTGAGATGAGATGTGTGACAGTCCAGTCGGAAGTCTCAGGCAGAGATAC	468				
QY 241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCCTGAGC	300				
Db 469	CACATGCCCAAGTGTATGTGGCTAAAGTCGCCCAAGGGTGGCTGTACGAGGGCCTGAGC	528				
QY 301	AGGGAAGAACAGAGAACTGCTGTCTTACCTGGGAACCTGGAGGGGCTTCTCATC	360				
Db 529	CGGGAGAAAGCCGAGGAAGTCTCTGTACTGTGGAAACCCCGAGGGGCTTCTCATC	588				
QY 361	CGGGAGACCCAGACCCAGAGAGGCTTTACTCTCTGTCAAGTCCGCGCTCAGCCGCCCTGCA	420				
Db 589	CGGGAGACCCAGACCCAGAGAGGCTTACTATTCCTGTCCGTCCGACTCAGCCGCCCTGCA	648				

QY	421	TCCTGGACCGGATCAGACACTACAGGATCCACTGCCTTGACAAATGGCTGCCTGTACATC	480
Db	649	TCTTGGACCGGATCAGACACTACAGGATACAGCTCTTGACAAATGGCTGCCTGTACATC	708
QY	481	TCACCGGCGCTCACTTCCCTCTACCTCCAGCGCCCTGGTGAGCCATTACTGTGAGCTGGCG	540
Db	709	ACACCTCGCCTCACTTCCCTCTACCTCCAGCGCTTGGTGAGCATTACTGTGAGCTAGCA	768
QY	541	GATGACATCTGCTGCCTACTCAAGGAGCCCTGTGTCTGCAGAGGGCTGGCCGCTCCCT	600
Db	769	GATGCATCTGCTGTCTCCCTCAGGGAGCCGTGTGTCTGCAGAAAGCTTGGGCCACTACCT	828
QY	601	GGCAAGATATACCCCTACCTGTGACTGTGCAGAGGACACCACTCAACTGGAAAGAGCTG	660
Db	829	GGCAAGATATACCTCCACCTGTGACTGTGCCAACAATCATCACTAAATGGAAAAAGCTG	888
QY	661	GACAGCTCCCTCCTGTCTTCTGAAG--CTGCCACAGGGAGGAGTCTCTTCTCAGTGAG	717
Db	889	GACCGAGCCTCCTGTCTTCTGAAG--CTGCCACAGGGAGGAGTCTCTTCTCAGTGAG	948
QY	718	GGTCTCCGGGAGTCCCTCTAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGGAT	777
Db	949	GGGCTCCGAGAGTCCCTCTAGCTTCTACATCAGCCTGGCTGAGGACCCCTTGGATGATGCT	1008
QY	778	GATGCCTAG 786	
Db	1009	TAGCCCTGG 1017	

RESULT 6	AL541041	LOCUS	DEFINITION
	AL541041	1032 bp	mRNA linear EST 24-MAR-2004
	AL541041	Homo sapiens	PLACENTA Homo sapiens cDNA clone CS0DE005YK23
	5-PRIME,	mRNA sequence.	

ACCESSION	AL541041	
VERSION	AL541041.3	GI:45716635
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

Chordata; Euteleostomi; Eukaryota; Metazoa; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1032)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:30544829.

Contact: Genoscope
Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

This sequence belongs to sequence cluster 9825.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0DE05AF12QP1&c=9825.r>.

FEATURES	Location/Qualifiers
source	1. .1032

```

Location/Qualifiers
1. . 1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YK23"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched, into
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

ORIGIN

Query Match	67.3%;	Score 529;	DB 1;	length 1032;
Best Local Similarity	96.5%;	Pred. No. 1.1e-128;		
Matches 526;	Conservative 12;	Mismatches 7;	Indels 0;	Gaps 0;

QY	1	ATGGAAGTCTGCCAGCAGAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	60
Db	395	ATGGGAAGTCTGCCAGCAGAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	454
QY	61	CAAGGCCAGGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTG	120
Db	455	CAAGGCCAGGGACCTGTGACCATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTG	514
QY	121	GGCAGTTTCCCGGCAGGTGGCCCCGCCGAGCTGTGCTGAGACTCCGGGAGCCATTGACC	180
Db	515	GGCAGTTTCCCGGCAGGTGGCCCCGCCGAGCTGTGCTGAGACTCCGGGAGCCATTGACC	574
QY	181	ATCGTCTCTGAGGATGGAAGTGGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT	240
Db	575	ATCGTCTCTGAGGATGGAAGTGGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT	634
QY	241	AACATCCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	300
Db	635	AACATCCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAGC	694
QY	301	AGGGAGAAAGCAGAGGAAGTCTGTGTTACCTGGGAACCTGGAGGGCCTTCCTCATC	360
Db	695	AGGGAGAAAGCAGAGGAAGTCTGTGTTACCTGGGAACCTGGAGGGCCTTCCTCATC	754
QY	361	CGGGAGAGCCAGAGCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCCTCAGCCGCCCTGCA	420
Db	755	CGGGAGAGCCAGAGCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCCTCAGCCGCCCTGCA	814
QY	421	TCCCTGGGACCCGATCAGACACTACAGGATCCACTGCCTTGACAAATGGCTGCTGATAC	480
Db	815	TCCCTGGGACCCGATCAGACACTACAGGATCCACTGCCTTGACAAATGGCTGCTGATAC	874
QY	481	TCAACCGGCTCACCTTCCCTCACTCCAGGCCCTGGTGACCAATTACTGAGCTGGCG	540
Db	875	TCAACCGGCTTAACTTCCCTCACTCCAGGCCCTGGTGACCAATTACTTAACTGAGGGG	934
QY	541	GATGA 545	
Db	935	RTTAA 939	

RESULT 7	AK020837	LOCUS	DEFINITION
AK020837	926 bp	mRNA	linear
Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone: A930009E21 product: MODULATOR OF ANTIGEN RECEPTOR SIGNALING			
MARS, full insert sequence.			

ACCESSION	AK020837	GI:12861542
VERSION	AK020837.1	
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE	1
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED
2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Genome Res. 10 (10), 1617-1630 (2000)	20499374	11042159

REFERENCE
AUTHORS
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ieshi,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
JOURNAL
MEDLINE
PUBMED
11076861
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL
REFERENCE
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 926)
JOURNAL
REFERENCE
AUTHORS
6
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
please visit our web site (http://genome.gsc.riken.jp/) for further
details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through two rounds of normalization to Rot = 20.0 and subtraction
to Rot = 458.8. Second strand cDNA was prepared with the primer
adapter of sequence [5'
GAGAGAGAGATTCGAGTTAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
bulk excision from lambda FLIC 1. Cloning sites, 5' end: SalI; 3'
end: BamHI. Host: DH10B. Retina RNA was provided by Stefano
Gustincich (Department of Neurobiology, Harvard Medical School, 220
Longwood Ave., Boston, MA 02115, USA) whose assistance is
gratefully acknowledged.
Retina RNA was provided by Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged.
FEATURES
source
1.926
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

ORIGIN
polyA_signal
polyA_site
Query Match 60.8%; Score 478; DB 3; Length 926;
Best local similarity 83.1%; Pred. No. 3.3e-115;
Matches 557; Conservative 0; Mismatches 110; Indels 3; Gaps 1;
CDS
/db_xref="FANTOM DB:A930009E21"
/db_xref="taxon:10090"
/clone="A930009E21"
/tissue_type="retina"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
128..667
/note="unnamed protein product; MODULATOR OF ANTIGEN
RECEPTOR SIGNALING MARS (SPTR) AAL38196, evidence: FASTY,
100%ID, 100%length, match=777)
putative"
/codon_start=1
/protein_id="BAB32223.1"
/db_xref="GI:12861543"
/translation="MPSVYVAKVAHGWLYEGLSREKAEELLPLGNPGAFILRSQT
RRGYSLSVRLSRPASWDRIHRYRIQRLDNGWLYISPLTFPSLHVLVEHYSELADGI
CCPLREPCVLQKLGPLPGKDTPPVTVPPTSLSMVKLDRLFLLEAPASGEASLISEG
LRESLSYSISLADPLDDA"
907..912
/note="putative"
926
/note="putative"
polyA_signal
polyA_site
Query Match 60.8%; Score 478; DB 3; Length 926;
Best local similarity 83.1%; Pred. No. 3.3e-115;
Matches 557; Conservative 0; Mismatches 110; Indels 3; Gaps 1;
CDS
/db_xref="FANTOM DB:A930009E21"
/db_xref="taxon:10090"
/clone="A930009E21"
/tissue_type="retina"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
128..667
/note="unnamed protein product; MODULATOR OF ANTIGEN
RECEPTOR SIGNALING MARS (SPTR) AAL38196, evidence: FASTY,
100%ID, 100%length, match=777)
putative"
/codon_start=1
/protein_id="BAB32223.1"
/db_xref="GI:12861543"
/translation="MPSVYVAKVAHGWLYEGLSREKAEELLPLGNPGAFILRSQT
RRGYSLSVRLSRPASWDRIHRYRIQRLDNGWLYISPLTFPSLHVLVEHYSELADGI
CCPLREPCVLQKLGPLPGKDTPPVTVPPTSLSMVKLDRLFLLEAPASGEASLISEG
LRESLSYSISLADPLDDA"
907..912
/note="putative"
926
/note="putative"

Db 664 TTAGCCCTGG 673

RESULT 8
BY742155
LOCUS
DEFINITION BY742155 660 bp mRNA linear EST 17-DEC-2002
BY742155 RIKEN full-length enriched, adult retina Mus musculus cDNA
clone A930009E21 5', mRNA sequence.

ACCESSION BY742155
VERSION
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasaki, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
JOURNAL MEDLINE
PUBMED
COMMENT

12466851
22354683
Nature 420, 563-573 (2002)

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES
source

Location/Qualifiers
1. 660
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A930009E21"
/tissue_type="retina"
/dev_stage="adult"
/lab_host="MDH10B"
/clone_1lb="RIKEN full-length enriched, adult retina"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5'-GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
Phi I. ~Retina RNA was provided by Stefano Gustincich,
Department of Neurobiology, Harvard Medical School, 220
Longwood Ave., Boston, MA02115, USA, whose assistance we
gratefully acknowledge."

ORIGIN

Query Match 57.7%; Score 453.8; DB 6; Length 660;
Best Local Similarity 83.6%; Pred. No. 7.7e-109;
Matches 539; Conservative 0; Mismatches 102; Indels 4; Gaps 2;

QY 120 GGGCAGTTTCCCGCAGAGTGGCCCGCCGAGCTGTGCGTGAGACTCGGGAGCCATTGAC 179
Db 4 GGGCAGTTTCCCGCAGAGTGGCCCGCCGAGCTGTGCGTGAGACTCGGGAGCCATTGAC 63
QY 180 CATCGTCTGTGAGGATGAGACTGGTGGACCGGTGCTGTGTAAGTTCAGGCAGAGAGTA 239
Db 64 CATCATCTGTGAGGATGAGACTGGTGGACCGGTGCTGTGTAAGTTCAGGCAGAGAGTA 123
QY 240 TAACATCCCGCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCTGAG 299
Db 124 CCACATGCCAGTGTGTATGTGGCTAAAGTCCGCCACGGGTGGCTGTATGAGGGCCTGAG 183
QY 300 CAGGAGAAAGCAGAGGAAGTGTGTTTACCTGAGAACCTTGAGGGGCTTCTCAT 359
Db 184 CCGGAGAAAGCAGAGGAAGTGTGTTTACCTGAGAACCTTGAGGGGCTTCTCAT 243
QY 360 CCGGAGAGCCAGACAGAGAGAGGCTTACTCTGTGTCAGTCCGCTCAGCCGCTGC 419
Db 244 CCGGAGAGCCAGACAGAGAGAGGCTTACTCTGTGTCAGTCCGCTCAGCCGCTGC 303
QY 420 ATCTGGGACCGGATCAGACACTACAGGATCAGCTGCTTGACATGGCTGGTATACAT 479
Db 304 ATCTGGGACCGGATCAGACACTACAGGATCAGCTGCTTGACATGGCTGGTATACAT 363

QY 480 CTCACCGCCCTCACCTTCCCTCACTCCAGGCCCTGTGGACATTACTGTAGCTGGC 539
| | | | |
Db 364 CTCACCTCGCTCACCTTCCCTCACTCCAGCCTGTGGACATTACTGTAGCTAGC 423
| | | | |
QY 540 GGATGACATCTGCTGCTCAAGGAGCCCTGTGCTCTGAGAGGGCTGGCCGCTCC 599
| | | | |
Db 424 AGATGGCATCTGCTGCTCCCTCAGGAGCGGTGTCTCTGAGAGCTTGGCCACTACC 483
| | | | |
QY 600 TGGCAAGATATACCCCTACCTGTGCTGAGAGAGACACCACTCACTGAAAGAGCT 659
| | | | |
Db 484 TGGCAAGATATACACCTCCACCTGTGCTGACACATCATCTAATTGAAAAAGCT 543
| | | | |
QY 660 GGACAGCTCCCTCCTGTTTCTGAG--CTGCCACAGGGGAGAGTCTCTTCTCAGTGA 716
| | | | |
Db 544 GGACCGTAGCCTCCTGTTTCTGAAAGACCTGCAAGTGGGAGGCATCTCTCATTTGA 603
| | | | |
QY 717 GGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGA 761
| | | | |
Db 604 GGGGCTCCAGAGT-CCTCACCTTCTACCTCATCCCGTTGATGA 647
| | | | |

RESULT 9

BG284179 566 bp mRNA linear EST 21-FEB-2001
LOCUS 602408226F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520382 5',
DEFINITION mRNA sequence.

ACCESSION BG284179

VERSION BG284179.1 GI:13034866

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10418 row: C column: 07
High quality sequence start: 2
High quality sequence stop: 566.

FEATURES

source

1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4520382"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/note="Organ: prostate; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 51.5%; Score 405; DB 4; Length 566;
Best Local Similarity 100.0%; Pred. No. 6e-96;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 GGCTCTTACTCTGTGTCAGTCCGCTCAGCGCCCTGCATCTGGACCGATCAGACAC 441
| | | | |
Db 13 GGCTCTTACTCTGTGTCAGTCCGCTCAGCGCCCTGCATCTGGACCGATCAGACAC 72
| | | | |

QY 442 TACAGATCCACTGCTTGCACATGGCTGGCTGTACATCTACCGGCTCACCTTCCC 501
| | | | |
Db 73 TACAGATCCACTGCTTGCACATGGCTGGCTGTACATCTACCGGCTCACCTTCCC 132
| | | | |
QY 502 TCACTCCAGGCCCTGTGTGACCATTACTGTAGCTGGCGGATGACATCTGCTACTC 561
| | | | |
Db 133 TCACTCCAGGCCCTGTGTGACCATTACTGTAGCTGGCGGATGACATCTGCTACTC 192
| | | | |
QY 562 AAGAGCCCTGTGTCTCTGACAGGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACC 621
| | | | |
Db 193 AAGAGCCCTGTGTCTCTGACAGGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACC 252
| | | | |
QY 622 GTGACTGTGACAGAGACACCACTCACTGAAAGAGCTGACAGCTCCCTGTTTCT 681
| | | | |
Db 253 GTGACTGTGACAGAGACACCACTCACTGAAAGAGCTGACAGCTCCCTGTTTCT 312
| | | | |
QY 682 GAAGCTGCCACAGGGGAGAGTCTTCTCTCAGTGAGGGTCCGGAGTCCCTCAGCTTC 741
| | | | |
Db 313 GAAGCTGCCACAGGGGAGAGTCTTCTCTCAGTGAGGGTCCGGAGTCCCTCAGCTTC 372
| | | | |

RESULT 10

BQ054265 986 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT_6830248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339
DEFINITION 5', mRNA sequence.

ACCESSION BQ054265

VERSION BQ054265.1 GI:19813605

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2125 row: i column: 12
High quality sequence stop: 515.

FEATURES

source

1..986
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936339"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using Zap-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 49.1%; Score 386.2; DB 5; Length 986;
Best Local Similarity 98.0%; Pred. No. 6.3e-91;

```
Matches 444; Conservative 0; Mismatches 3; Indels 6; Gaps 5;
QY 1 ATGGGAAGTCTGCCACAGAGAAGAAATCTTCCCAAGCCCAAGCTTGAGTTCTCTGTC 60
    |||
Db 279 ATGGGAAGTCTGCCACAGAGAAGAAATCTTCCCAAGCCCAAGCTTGAGTTCTCTGTC 338
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 120
    |||
Db 339 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 398
QY 121 GGCAGTTTCCCGCAGGTGGCCCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
    |||
Db 399 GGCAGTTTCCCGCAGGTGGCCCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 458
QY 181 ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT 240
    |||
Db 459 ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT 518
QY 241 AACATCCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300
    |||
Db 519 AACATCCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 578
QY 301 AGCGAGAAAGCAGAGAACTGCTGTTGTTACTTGGAAACCTTGAGGGG-CCTTCTCAT 359
    |||
Db 579 AGCGAGAAAGCAGAGAACTGCTGTTGTTACTTGGAAACCTTGAGGGG-CCTTCTCAT 638
QY 360 CCGGGAG-AGCCAGACCAGAGAGGCTCTTACTCTGTGAGT-CCGCTCAGCCG-CCC 416
    |||
Db 639 CCGGGAGAGAGCCAGACCAGAGAGGCTCTTACTCTGTGAGTCCCGCTCAGCCGCCCC 698
QY 417 TGCATCTCTGGG--ACCGATCAGACACTACAGG 447
    |||
Db 699 TGCATCTCTGGGACCGGATCAGACCTACAGG 731
```

```
RESULT 11
BQ053486 878 bp mRNA linear EST 29-MAR-2002
LOCUS BQ053486
DEFINITION AGENCOURT 6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253
5', mRNA sequence.
```

```
ACCESSION BQ053486
VERSION BQ053486.1 GI:19812826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NIH-MGC <http://mgc.nci.nih.gov/>.

Unpublished (1999) National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM2122 row: 1 column: 06

High quality sequence stop: 394.

Location/Qualifiers

1. .878

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5935253"

/tissue_type="natural killer cells, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_106"

/note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN
Query Match 42.1%; Score 331.2; DB 5; Length 878;
Best Local Similarity 93.1%; Pred. No. 2.1e-76;
Matches 390; Conservative 0; Mismatches 24; Indels 5; Gaps 4;

```
QY 1 ATGGGAAGTCTGCCACAGAGAAGAAATCTTCCCAAGCCCAAGCTTGAGTTCTCTGTC 60
    |||
Db 273 ATGGGAAGTCTGCCACAGAGAAGAAATCTTCCCAAGCCCAAGCTTGAGTTCTCTGTC 332
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 120
    |||
Db 333 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 392
QY 121 GGCAGTTTCCCGCAGGTGGCCCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
    |||
Db 393 GGCAGTTTCCCGCAGGTGGCCCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 452
QY 181 ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT 240
    |||
Db 453 ATCGTCTCTGAGGATGAGACTGTGTGACCGGTGCTGTGAAGTCTCAGGCAGAGAGTAT 512
QY 241 AACATCCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300
    |||
Db 513 AACATCCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 572
QY 301 AGCGAGAAAGCAGAGAACTGCTGTTGTTACTTGGAAACCTTGAGGGG-CCTTCTCAT 359
    |||
Db 573 A--GAAAAAACCAAGACTGCTGTTGTTACTTGGAAACCTTGAGGGG-CCTTCTCTC 630
QY 360 CCGGAGAGCCAGACCAGAGAGGCT-CTTACTCTCTGT-CAGTCCGCTCAGCCGCCCC 416
    |||
Db 631 CCGGAGAGACCAGACCAGAGAGGCTCTTACTCTCTGTGTTAGTCCGCTTAGCCGCCCC 689
```

```
RESULT 12
BG178487 778 bp mRNA linear EST 06-FEB-2001
LOCUS BG178487
DEFINITION 602328305F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429896 5',
mRNA sequence.
```

```
ACCESSION BG178487
VERSION BG178487
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NIH-MGC <http://mgc.nci.nih.gov/>.

Unpublished (1999) National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM10182 row: 1 column: 01

High quality sequence stop: 657.

Location/Qualifiers

1. .778

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"


```

/clone="IMAGE:4429896"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/notes="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

ORIGIN

```

Query Match      41.6%; Score 326.8; DB 4; Length 778;
Best Local Similarity 94.6%; Pred. No. 3e-75;
Matches 371; Conservative 0; Mismatches 17; Indels 4; Gaps 3;

```

```

QY 1 AATGGAGTCTGCCCCAGACAGAAATAATCTTCCCAAGCCCAAGCTTGAGTTCTCTGTC 60
Db 362 AATGGAGTCTGCCCCAGACAGAAATAATCTTCCCAAGCCCAAGCTTGAGTTCTCTGTC 421
QY 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 422 CAAGGCCA-GGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 121 GGCAAGTTTCCCGGAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
Db 481 GGCAAGTTTCCCGGAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 540
QY 181 ATGCTCTGAGGATGAGACTGTGTGACGCGTGTCTGTAAGTCTCAGGACAGAGATAT 240
Db 541 ATGCTCTGAGGATGAGACTGTGTGACGCGTGTCTGTAAGTCTCAGGACAGAGATAT 600
QY 241 AACATCCCCCAGCGTCCACGTGGCCCAAGTCTCCCATGGGTGGCTGTATAGAGGCTGAGC 300
Db 601 AACATCCCCCAGCGTCCACGTGGCCCAAGTCTCCCATGGGTGGCTGTATAGAGGCTGAGC 660
QY 301 AGGAGAG-AAGCAGAGAGACTGTGTGTACCTGGGAACCCCTGAGAGGCGCTTCC--TC 357
Db 661 AGGAGAGACAGCAGAGAGACTGTGTGTACCTGGGAACCCCTGAGAGGCGCTTCC--TC 420
QY 358 ATCCGGAGAGAGCCAGACAGAGAGAGGCTCTTA 389
Db 721 TCCCGGAGAGAGCCAGACAGAGAGAGGCTCTTA 752

```

```

RESULT 13
AL844311 597 bp mRNA linear EST 30-JUL-2002
LOCUS AL844311 pool YT_1ib_v_SPD Homo sapiens cDNA, mRNA sequence.
DEFINITION AL844311
ACCESSION AL844311
VERSION AL844311.1 GI:22019093
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 597)
Ashcroft,K., Bechel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
Sheridan,E.
REFERENCE Homo sapiens EST sequence
AUTHORS Unpublished (2002)
JOURNAL Contact: The Sanger Centre
COMMENT The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name: sccdl0818.154136A
Homo sapiens EST sequence. This sequence was generated as part of
The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library
resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool YT_1ib v SPD cDNA library. Further information can be found at
http://www.sanger.ac.uk/Teams/Team69/.

```

FEATURES

source

```

Location/Qualifiers
1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool YT_1ib_v_SPD"

```

ORIGIN

```

Query Match      41.0%; Score 322.4; DB 1; Length 597;
Best Local Similarity 99.7%; Pred. No. 4.2e-74;
Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 463 AATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGAC 522
Db 13 AATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGAC 72
QY 523 CATTAAGTCTGAGTGGCGGATGACATCTGTGCTGCTCAAGAGAGAGAGAGAGAGAGAGAG 582
Db 73 CATTAAGTCTGAGTGGCGGATGACATCTGTGCTGCTCAAGAGAGAGAGAGAGAGAGAGAG 132
QY 583 AGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACTGTGACTGTGACAGAGACCA 642
Db 133 AGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACTGTGACTGTGACAGAGACCA 192
QY 643 CTCAACTGAAAGAGCTGACAGAGCTCCCTCTGTTTCTGAAGCTGCCACAGAGGAGAG 702
Db 193 CTCAACTGAAAGAGCTGACAGAGCTCCCTCTGTTTCTGAAGCTGCCACAGAGGAGAG 252
QY 703 TCTCTTCTCAGTGGGTCTCCGGAGTCCCTCACTCTCAATCAGAGGAGAGAGAGAGAG 762
Db 253 TCTCTTCTCAGTGGGTCTCCGGAGTCCCTCACTCTCAATCAGAGGAGAGAGAGAGAGAG 312
QY 763 GCTGTCTCTTTGGATGATGCTTAG 786
Db 313 GCTGTCTCTTTGGATGATGCTTAG 336

```

```

RESULT 14
BU944126 960 bp mRNA linear EST 18-OCT-2002
LOCUS BU944126
DEFINITION AGENCOURT 10545003 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6728350 5', mRNA sequence.
ACCESSION BU944126
VERSION BU944126
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 960)
NIH-MGC http://mgi.nci.nih.gov/.
REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-romail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3049 row: m column: 21
High quality sequence stop: 628.
location/Qualifiers
1..960
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6728350"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

FEATURES

source

```

location/Qualifiers
1..960
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6728350"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

ORIGIN

Query Match 40.8%; Score 321; DB 5; Length 960;
Best Local Similarity 88.4%; Pred. No. 1.1e-73;
Matches 381; Conservative 0; Mismatches 0; Indels 50; Gaps 1;
QY 356 TCATCCGGAGAGACCCAGACAGAGAGGCTTACTCTCTGTCAGTCCGCGCTCAGCCGCC 415
Db 1 TCATCCGGAGAGACCCAGACAGAGAGGCTTACTCTCTGTCAGTCCGCGCTCAGCCGCC 60
QY 416 CTGCATCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGCTGT 475
Db 61 CTGCATCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACAATGGCTGCTGT 120
QY 476 ACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTTGAGC 535
Db 121 ACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTTCT 176
QY 536 TGGCGGATGACATCTGTGCTGCTCAAGAGGCCCTGTGTCTGACAGGGGCTGCGCCGC 595
Db 177 -----GAGGCTGCGCCGC 190
QY 596 TCCCTGGCAAGATATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGGAAG 655
Db 191 TCCCTGGCAAGATATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGGAAG 250
QY 656 AGCTGGACAGACTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTTCTCTCAGTG 715
Db 251 AGCTGGACAGACTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTTCTCTCAGTG 310
QY 716 AGGTTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGG 775
Db 311 AGGTTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTGG 370
QY 776 ATGATGCCTAG 786
Db 371 ATGATGCCTAG 381

RESULT 15
AL844307 614 bp mRNA linear EST 30-JUL-2002
LOCUS AL844307 pool_AK_lib_v_SPD Homo sapiens cDNA, mRNA sequence.
DEFINITION AL844307
ACCESSION AL844307
VERSION AL844307.1 GI:22019089
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 614)
AUTHORS Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
Sheridan,E.
TITLE Homo sapiens EST sequence
JOURNAL Unpublished (2002)
COMMENT Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
Email: humquery@sanger.ac.uk
Sanger Centre name : sccdl0816.154136A

The Wellcome Trust Sanger Institute program to identify and
annotate genes in the human genome. Incomplete or unconfirmed genes
are experimentally analysed using a variety of cDNA library

resources. This sequence was obtained from a PCR product generated
from a pool of up to 100,000 cDNA clones derived from
pool_AK_lib v SPD cDNA library. Further information can be found at
<http://www.sanger.ac.uk/Teams/Team69/>.

FEATURES

source

1. 614
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="20"
/clone_lib="pool_AK_lib_v_SPD"

ORIGIN

Query Match 40.4%; Score 317.4; DB 1; Length 614;
Best Local Similarity 99.7%; Pred. No. 8.8e-73;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 468 CTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTA 527
Db 1 CTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTA 60
QY 528 CTCTGAGCTGGCGGATGACATCTGCTGCTTACTCAAGAGCCCTGTGTCTGACAGGGC 587
Db 61 CTCTGAGCTGGCGGATGACATCTGCTGCTTACTCAAGAGCCCTGTGTCTGACAGGGC 120
QY 588 TGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACCACTCAA 647
Db 121 TGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGACAGAGACCACTCAA 180
QY 648 CTGGAAGAAGCTGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCT 707
Db 181 CTGGAAGAAGCTGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTCT 240
QY 708 TCTCAGTGAAGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGT 767
Db 241 TCTCAGTGAAGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGT 300
QY 768 CTCTTTGATGATGCCCTAG 786
Db 301 CTCTTTGATGATGCCCTAG 319

Search completed: November 16, 2004, 23:30:37
Job time : 3256 secs